4

AD-A212 520

AD

Technical Memorandum 5-89

THE EFFECTS OF NOISE ON SPEECH AND WARNING SIGNALS

Alice H. Suter
Gallaudet University

June 1989
AMCMS Code 611102.74A0011

Approved for public release;
distribution is unlimited.

This report was prepared by Gallaudet University for the U.S. Army Human Engineering Laboratory. This report has received a technical review by the Human Engineering Laboratory but it has not received an editorial review by the laboratory. This report is presented in the interest of providing information in a timely manner.

U. S. ARMY HUMAN ENGINEERING LABORATORY

Aberdeen Proving Ground, Maryland

DISCLAIMER NOTICE

THIS DOCUMENT IS BEST QUALITY PRACTICABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) Technical Memorandum 5-89
6a. NAME OF PERFORMING ORGANIZATION Gallaudet University 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Human Engineering Laboratory
6c. ADDRESS (City, State, and ZIP Code) Washington, DC 20002	7b. ADDRESS (City, State, and ZIP Code) Aberdeen Proving Ground, MD 21005-5001
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
6.11.02	1L161102B74A		

11. TITLE (Include Security Classification)
The Effects of Noise on Speech and Warning Signals

12. PERSONAL AUTHOR(S)
Suter, Alice H.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM ____ TO ____	1989, June	55

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP
23	02	
25	04	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
speech communication, speech interference, warning signals, masking, intelligibility

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

To assess the effects of noise on speech communication it is necessary to examine certain characteristics of the speech signal. Speech level can be measured by a variety of methods, none of which has yet been standardized, and it should be kept in mind that vocal effort increases with background noise level and with different types of activity. Noise and filtering commonly degrade the speech signal, especially as it is transmitted through communications systems. Intelligibility is also adversely affected by distance, reverberation, and monaural listening. Communication systems currently in use may cause strain and delays on the part of the listener, but there are many possibilities for improvement.

(see reverse side)

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Georges Garinther	22b. TELEPHONE (Include Area Code) (301) 278-5984 22c. OFFICE SYMBOL SLCHE-BR

19. (continued)

Individuals who need to communicate in noise may be subject to voice disorders. Shouted speech becomes progressively less intelligible at high voice levels, but improvements can be realized when talkers use "clear speech." Tolerable listening levels are lower for negative than for positive S/Ns, and comfortable listening levels should be at a S/N of at least 5 dB, and preferably above 10 dB.

Popular methods to predict speech intelligibility in noise include the Articulation Index ~~(AI)~~, Speech Interference Level ~~(SIL)~~, Speech Transmission Index ~~(STI)~~ and the sound level meter's A-weighting network. This report describes these methods, discussing certain advantages and disadvantages of each, and shows their interrelations. (KR)

Communication is considered "just reliable" at an AI of 0.3 to 0.45 although little evidence is available to support these criteria. Likewise, there is little information available on the specific types and amounts of communication needed for various operations, and the only available evidence on the consequences of degraded speech tends to be anecdotal or subjective.

Audible warning signals should be at least 15 dB but no more than 25 dB above masked threshold. Temporal, spectral, and ergonomic aspects should emphasize attention demand, relevance, and appropriate level of priority without being unduly aversive.

AMCMS Code 611102.74A0011 Technical Memorandum 5-89

THE EFFECTS OF NOISE ON SPEECH AND WARNING SIGNALS

Alice H. Suter
Gallaudet University
Washington, DC

June 1989

Accesion For	
NTIS CRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and / or Special
A-1	

APPROVED: [signature]
JOHN D. WEISZ
Director
Human Engineering Laboratory

This report was prepared by Gallaudet University for the U.S. Army Human Engineering Laboratory. This report has received a technical review by the Human Engineering Laboratory but it has not received an editorial review by the laboratory. This report is presented in the interest of providing information in a timely manner.

U.S. ARMY HUMAN ENGINEERING LABORATORY
Aberdeen Proving Ground, Maryland 21005-5001

CONTENTS

I. INTRODUCTION

The effective communication of speech and warning signals is vital to the success of a military program. The consequences of communication failures can range from a minor irritation to a major disaster, depending on the importance of the incorrectly perceived message. These communication failures can be costly in terms of mission objectives, equipment, and, in the extreme, human life. Adequate technology exists to permit effective communication in most situations, but it is not always implemented. In some conditions, a high level of intelligibility is unnecessary because the communication task is very simple. In others, however, highly intelligible communications are needed to convey complex or unexpected messages in emergency situations. It is important to assess each communication situation so that the right balance can be made between economy and program effectiveness. Unnecessary sophistication in communication systems should be avoided, but too much emphasis on economy can lead to greater expense in the long run.

The purpose of this literature search and analysis is: (1) to elucidate the present state of information on the effects of noise on the perception and recognition of speech and warning signals; (2) to describe some of the circumstances in which communication improvements or degradations may occur; and (3) to identify additional information collection or research projects that will improve speech and signal recognition in military environments.

To obtain an understanding of speech and warning signal communication in the military context, it is first necessary to explore some theoretical and practical aspects of communication, especially as it is affected by noise. The report will cover speech variables, namely speech level, materials used for testing communication systems, and distortions of speech by filtering and masking. It will include a discussion of the transmission of speech from talker to listener; various talker and listener variables, such as the effect of non-native languages on both; and some of the more prominent methods for predicting the effects of noise and other degrading factors on speech intelligibility. The report will conclude with discussions of criteria for acceptable communication and for warning signal detection, and a number of recommendations for future research.

II. SPEECH VARIABLES

The intelligibility of speech depends on a large number of variables. The framers of ANSI S3.14 (ASA, 1977) divide them into acoustic, non-acoustic, and random or quasirandom factors. Acoustic factors include the level and spectrum of the speech signal at the listener's ear; the level, spectral, and temporal characteristics of the interfering noise; differences in the spatial locations of the speech and noise sources; and reverberation effects. Non-acoustic factors include the talker's speech habits, the size of the message set, the probability of occurrence of each unit, the listener's motivation and familiarity with the speech material, and visual cues. Random or quasirandom

factors, which set an "upward bound" on the precision with which intelligibility can be estimated, include individual differences between talkers and listeners, day-to-day variations in their effectiveness, effects of randomization in the choice of test material, random sampling errors, and the listener's age and hearing sensitivity (ASA, 1977 p. 1).

A. Speech Level

Any predictions of speech intelligibility are likely to be influenced by the procedure used to measure speech level. One of the difficulties is the wide dynamic range of speech, which is as much as 30 dB between the most and least intense phonemes (Webster, 1984; Pearsons, 1983; Hood and Poole, 1977). Another is a satisfactory method of accounting for the pauses between utterances. Various measurement methods have been proposed. One of the most popular methods is the long-term rms level monitored with a sound level meter or a VU meter. However, this method involves a certain amount of subjective judgement, and, according to Pearsons (1983), the speech sample should be at least 10 seconds long. Kryter (1984) maintains that the average A-weighted peak level of each word measured with a sound level meter set on slow response is approximately equal to the unweighted L_{eq}. Pearsons (1983) believes that the integrating sound level meter or computer shows promise (see also Suter, 1978), but points out that there are no standard techniques available.

Standardization is currently being considered by Working Group S3-59 for ANSI S3.38, "Measurement of Speech Levels" (ASA, 1986). A preliminary draft of this standard favors a method called the Equivalent Peak Level (EPL) developed by Brady (1968), with long-term rms measured in real-time as an alternative. The EPL method consists of measuring the rms level above an arbitrary threshold and calculating the peak of a log-uniformly distributed speech sample that would have the same rms level. The advantages of EPL are that it is (1) expressed by a single number, (2) uninfluenced by silent intervals, (3) independent of the threshold setting of the speech detector, and (4) follows known level changes on a dB for dB basis (Brady, 1968).

Although the various methods identify different speech levels, the relationships between these levels are fairly uniform. Most investigations show the unweighted rms level to be about 4 dB above A-weighted rms level, and the EPL to be 8 to 10 dB above unweighted rms (Pearsons, 1983; Steeneken and Houtgast, 1978). According to Kryter (1962a), speech peaks, defined as the level exceeded by only 1% of the speech signal, are equal to the rms level +12 dB. The long-term rms level may be estimated by taking the average speech peaks in quiet, measured with a sound level meter set on C-weighting and slow response, and subtracting 3 dB (Kryter, 1962a).

Speech level will change according to the vocal effort expended. Pickett (1956) found the range of vocal force varied from 36 dB, the level of a forced whisper, to 90 dB for a heavy shout. Figure 1 shows speech level as a function of vocal effort according to Pearsons et al. (1977) and including data from Beranek (1954). The entire range extends from about 48 dB to 92 dB.

People will increase their vocal effort automatically with increasing distance between talker and listener and with elevation in background noise level. Gardner (1966) found that people raise their voices approximately 2 dB

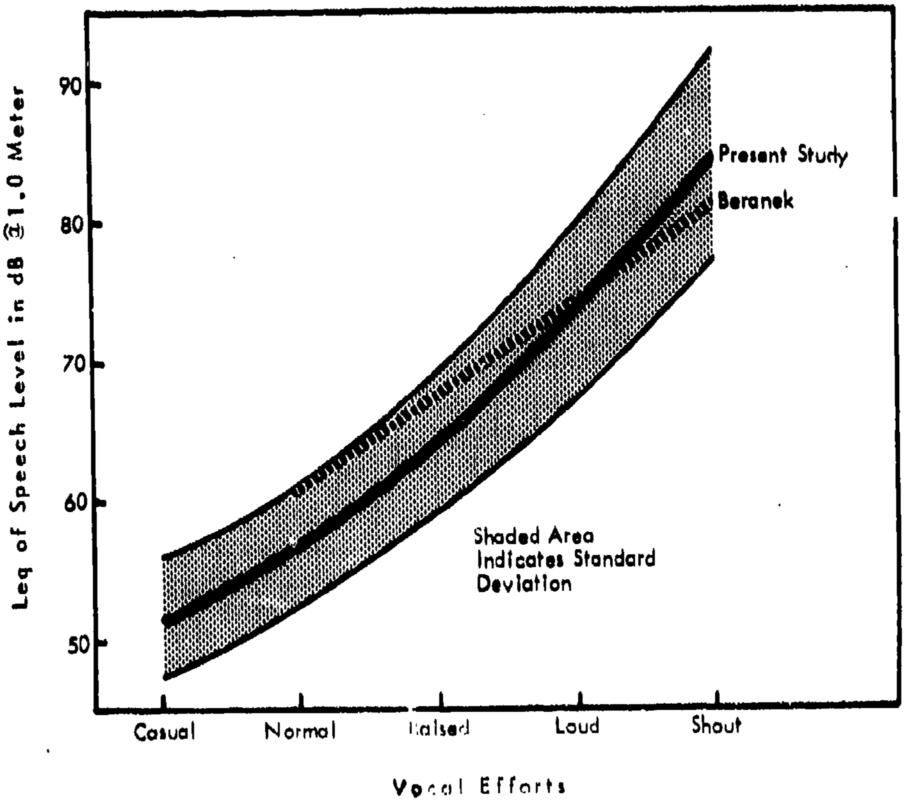

Figure 1. Speech levels for various vocal efforts.

Note. From Speech Levels in Various Noise Environments (EPA-600/1-77-025) by K. S. Pearsons, R. L. Bennett, and S. Fidell, 1977, U.S. Environmental Protection Agency. Reprinted by permission.

with every doubling of distance between about 1 and 4 meters. Kryter (1946) reports a 3-dB increase and Webster and Klumpp (1962) report a 5-dB increase for every 10-dB increase in background noise level. Webster and Klumpp (1962) identified the same increase in vocal effort as a result of a doubling in the number of talkers around a communicating pair (the "cocktail party" effect). Pearsons and his colleagues (1977) measured speech levels in face-to-face conversation at one meter, and found an increase of 6 dB for every 10 dB increase in background noise level between 48 and 70 dB, above which talker and listener moved closer together.

Changes in the speech spectrum and rate of utterance also occur as vocal effort increases. Webster and Klumpp (1962) found that speech rate decreased with increasing noise level, although it tended to increase with increasing numbers of competing talkers. Figure 2, also from Pearsons et al. (1977), displays the definite shift toward higher frequency speech energy with increasing vocal effort. People change their vocal effort according to their activity, even without increases in background noise level. They tend to talk louder when reading prepared text than they do in casual conversation. They also raise their voices when talking before an audience, on the telephone, and even in the presence of a microphone (Webster, 1984). On the basis of data from van Heusden et al. (1979), Houtgast advocates distinguishing between public and private communication, with the former being 9 dB louder than the latter (Houtgast, 1980). The intelligibility of amplified speech remains good up to sound pressure levels as high as 120 dB, but as soon as noise is introduced, even with a speech-to-noise ratio as favorable as 15 dB, intelligibility begins to drop off above a sound pressure level of 90 dB (Pollack and Pickett, 1958). Overloading the auditory system is presumably responsible. Unamplified speech is another matter. Intelligibility falls off abruptly above a speech level of 78 dB. However at speech levels below 55 dB, intelligibility falls off gently at first and then abruptly (Pickett, 1956).

B. Speech Materials

Spoken language contains numerous constraints, which add to its redundancy and make it easier to understand. This is indeed fortunate for the hearing-impaired individual and for any listener in a time of emergency. These constraints result from any language's grammatical structure, the context of the word or sentence, limitations in vocabulary size, the length of words, and the listener's familiarity with the speech material. The greater the constraints, the higher the speech intelligibility scores for the same speech-to-noise ratio. An example of this is the relative intelligibility of specialized vocabularies, such as the ones used by air traffic controllers. Frick and Sumby (1952) describe four steps of constraints in pilots' receipt of control tower messages: from an infinite set of possible messages one moves to a set of alphabetical sequences, then to a set of English sentences, to air language with its own particular grammar, and finally to the tower messages with their own procedural constraints. The estimated redundancy with respect to what could have been conveyed is 96%. The authors note that this degree of redundancy is very inefficient in terms of information transfer, but they point out that communication systems tend to be noisy, and the communication link between pilot and control tower has a low tolerance for error, so redundancy provides an important form of insurance.

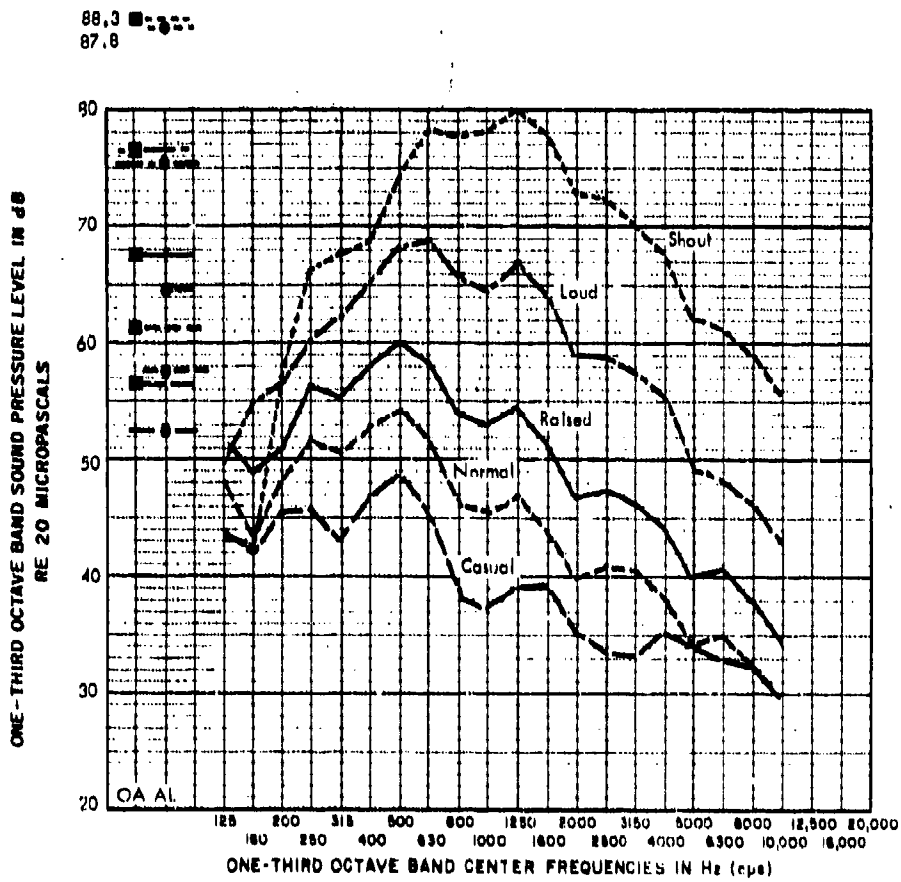

Figure 2. Average speech spectra for male talkers at five vocal efforts.

Note. From Speech Levels in Various Noise Environments (EPA-600/1-77-025) by K. S. Pearsons, R. L. Bennett, and S. Fidell, 1977, U.S. Environmental Protection Agency. Reprinted by permission.

Intelligibility increases directly as the number of possible words in a message set decreases. Similarly, for a given amount of intelligibility, the speech-to-noise ratio can be reduced with proportional decreases in the size of a message set. Miller et al. (1951) found that a decrease in message size from 256 to 4 monosyllables corresponded to a 12-dB decrease in speech-to-noise ratio. For this reason, "closed-set" tests, such as the Modified Rhyme Test (House et al., 1965), yield better intelligibility scores than "open set" tests of monosyllabic words or nonsense syllables, for a given speech-to-noise ratio. Other investigations have shown that long words are more intelligible than short ones (Rubenstein et al., 1959), and two-syllable words are more intelligible when the accent is on the second syllable (Black, 1952).

Figure 3 from ANSI S3.5 (1969) shows the relative intelligibility of various speech materials as a function of speech-to-noise ratio (represented by Articulation Index values). The order of difficulty is from the least intelligible, 1000 nonsense syllables; to 1000 phonetically balanced (PB) words; to rhyme tests, 256 PBs, and unfamiliar sentences; to familiar sentences; to the most intelligible, a vocabulary limited to 32 PB words. The committee cautions the reader that these relations are approximate, as they depend on the type of material and the skill of talkers and listeners.

Features within words can cause some words to be more intelligible (resistant to masking or filtering) than others. For example, prosodic features and vowels are more easily identified than consonants (Webster and Allen, 1972). Medial position phonemes are more intelligible than consonants in the initial and final position and final consonants are more easily identified than initial consonants under adverse conditions (Clarke, 1965).

In an effort to improve the reliability of the Harvard list of PBs, Hood and Poole (1977) noted that the intrinsic intelligibility of these words covered a range of at least 30 dB. (The authors considered this large range a necessary feature of a good intelligibility test.) By eliminating 5 "rogue lists", Hood and Poole brought the performance-intensity functions of the remaining 15 lists into close agreement. During this process, they analyzed the difficulty of all words in the 20 lists, having tested each word 36 times. The result is a table, which lists the relative intelligibility of all of the Harvard PBs, from most intelligible (jam, our, rope, wild, and will) to least intelligible (rave, fin, pun, and sup). This table could be useful in assessing the difficulty of words to be used in special phraseologies or for testing the articulation of specific systems.

The helpful redundancy in speech is derived from a number of different features, as explained above. In other words, it is as if we say the same thing in a variety of ways. The question arises, then, as to whether simple repetition of the same word will increase its intelligibility. Investigations of this question have produced moderately encouraging results. Miller et al. (1951) found that three successive presentations of the same word improved intelligibility by 5 to 10%, depending on the speech-to-noise ratio. Lazarus (1983) quotes a German colleague (Platte, 1978 and 1979) as finding that large variances can be avoided by triple repetition. Using the Harvard PBs, Thwing (1956) tested the effects of one through four presentations of the same word (e.g., "Item 26: dog, dog, dog") at three speech-to-noise ratios. The results showed a slight improvement between the first and second presenta-

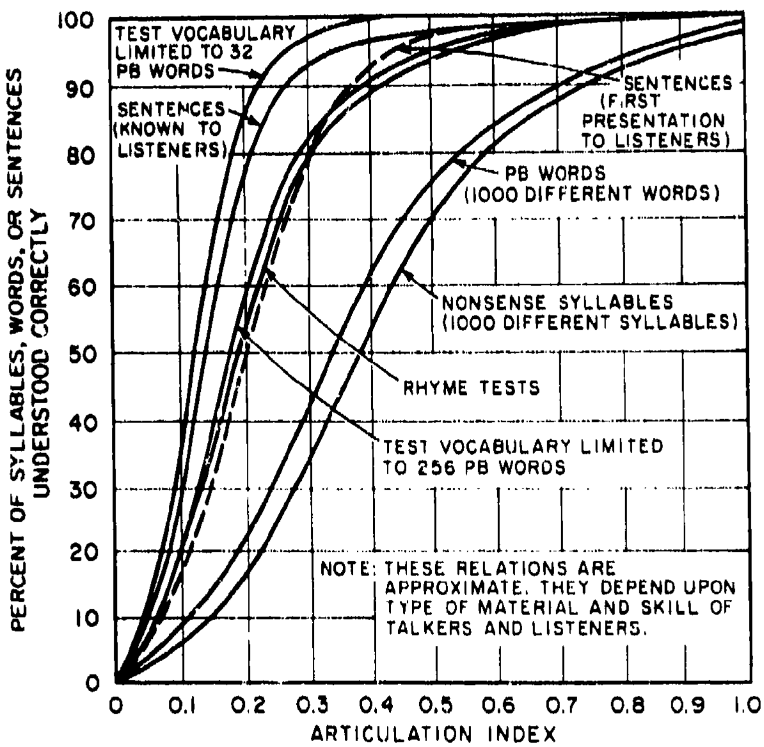

Figure 3. Relative intelligibility of various speech materials as a function of Articulation Index value.

tions, but nothing after that. The greatest improvement was at the most favorable speech-to-noise ratio. Other investigators found no improvement for repetition of numbers (Moser et al., 1954) or nonsense syllables (Black, 1955). Hood and Poole (1977) noticed that words duplicated (by chance) in separate lists were missed on some occasions but not on others. They cite Brandy (1966) as finding the same result, and suggest that the cause lies in slight variations in the talker's voice production, not only among different talkers but at different times with the same talker. So, at least for words, there appears to be a moderately beneficial effect of at least one repetition. In view of the increased opportunity for talker variations, it would seem reasonable that these benefits would be somewhat greater for phrases and short sentences.

As stated above, word familiarity is another important consideration in the intelligibility of a spoken message. According to Rubenstein and Pollack (1963), intelligibility is a simple power function of the probability of a word's occurrence. In an effort to develop word lists with familiarity greater than the Harvard PBs, Hirsh et al. (1952) developed the CID W-22 list of 200 familiar PBs, Peterson and Lehiste (1962) developed a CNC (Consonant Vowel Nucleus Consonant) list of 500 PBs, and Tillman and Carhart (1966) compiled the 200 words that comprise the NU Auditory Test 6, which was developed for and used extensively by the U.S. Air Force (Webster, 1972).

In a comprehensive compendium of speech testing materials, Webster (1972) discusses and reprints various speech materials and standard phraseologies used in testing communication systems. These include a selected list of Navy Brevity Code words, along with ICAO phonetic spelling words and digit pronunciation (Moser and Dreher, 1955), a transcription of radio transmissions of U.S. Naval aircraft over Vietnam (Webster and Allen, 1972), and a list of words frequently used in USAF aircraft compiled by Donald Gasaway. Gasaway's list includes statistics on word familiarity according to word-frequency counts from Thorndike and Lorge (1952), and a code showing whether they are represented in various standard word lists and among Brevity Code words. Webster's compendium also includes lists of tactical field messages from the U.S. Army Test and Evaluation Command (1971), 150 phrases from the flight deck of aircraft carriers developed by Klumpp and Webster (1960), and lists of aviation maintenance/supply support messages developed by Webster and Henry (NAVSHIPS, 1972).

One of the difficulties involved in speech testing using large sets of monosyllables, such as 1000 Harvard PBs, is the fact that talker and listener crews must be thoroughly trained. Webster (1972) states that such training takes weeks to perform! In an effort to reduce or eliminate training time, Fairbanks and his colleagues developed the closed-set Rhyme Test (Fairbanks, 1958), which has gone through a series of modifications (House et al., 1965; Kreul et al., 1968). An interesting innovation is the Tri-Word MRT (Williams, et al., 1976), where words are presented in triplets instead of individually. The principal advantage of this test is its speed: the investigators found that 51 words could be presented in only 2.3 minutes as opposed to 5 minutes for the MRT. Another variation developed by Voiers (1967) is the Diagnostic Rhyme Test (DRT), which can be used to identify the particular features of speech (in initial consonants only) that are affected by a communication system.

The American National Standard Method for Measurement of Monosyllabic Word Intelligibility (ANSI 3.2-1960 (R1982)), specifies the Harvard PBs as test materials. A current draft revision (ASA, 1988) has added the MRT and DRT. According to the new standard, the three tests have been shown to be highly correlated with each other as well as with other intelligibility test materials. All three tests will provide the same rank orders and magnitude of differences among systems when used with a large number of communication systems. The new draft also specifies the method outlined in ANSI S3.38 for measuring speech level (ASA, 1986). The scope of the new draft standard covers the testing of all kinds of communication systems (with the exception of speech recognition devices), including speech transmitted through air in rooms or out of doors; through telephonic systems including telephones, public address systems, and radios; or through complex environments including equipment, air, wire, fiber, radio, and water paths.

Sentence material can also be useful for testing communication systems. In addition to the original Harvard sentences (Hudgins _et al._, 1947), the CID sentences (Silverman and Hirsh, 1955) were developed to resemble "everyday" speech, and these sentences were modified to achieve homogeneity of sentence length to form the RCID sentences (Harris _et al._, 1961). Speaks and Jerger (1965) and Jerger _et al._ (1968) have developed synthetic sentences to reduce predictability of the key words. More recently, Kalikow _et al._ (1977) invented the SPIN test (Speech Intelligibility in Noise), which consists of two types of English sentences in speech-babble noise: one for which the key word is somewhat predictable from the context, and the other for which the key word cannot be predicted from the context. Both types of sentences are balanced for intelligibility, key-word familiarity and predictability, phonetic content, and length. Although its major application is in testing hearing-impaired persons, it has other uses, such as the evaluation of speech processing devices (Kalikow _et al._, 1977). The test has recently been revised by Bilger (1984), to achieve greater equivalence among test forms.

The choice of speech materials depends on many factors, including the availability of listeners and training time, the type of system, and the conditions of use. Webster (1978) points out that the midrange of a steep performance-intensity function is necessary for the best testing. Webster suggests that in very noisy conditions (AI of about 0.2), a closed-set test of rhyme words will yield about 50% intelligibility. At an AI of 0.35, an open set of 1000 PB words would be more appropriate because rhyme tests would yield about 85%, which would be at or above the "knee" of the function.* At AIs as high as 0.8, even 1000 nonsense syllables would produce intelligibility scores greater than 90%, so Webster advocates using other measures, such as reaction times, competing messages, or quality judgements (Webster, 1978).

---

*In his discussion of matching intelligibility tests to AI levels, Webster refers to an AI of 0.35 as corresponding to rhyme-test scores of 75%. However, the graph reprinted from ANSI S3.5, 1965 as Fig. 3 indicates rhyme scores of about 85% at this AI. This discrepancy may point up the caveat of the standard formulators, that these relations are approximate, and that they depend upon type of material and skill of talkers and listeners. It may also indicate the need for reexamining the interrelationships of these materials, especially in view of more recent additions to the available battery of test materials.

C. Distortions

According to Harris (1965), "...not more than half the time in everyday life do we listen to clearly enunciated speech in quiet." (p. 825). Distortions of speech, such as filtering and noise masking, are prevalent in all kinds of occupational environments and are common to many military situations, from offices and computer rooms to tracked vehicles and helicopters.

Filtering of speech occurs when it is passed through almost any transmission system, such as a telephone or a radio communication system. High-frequency speech sounds are most readily affected, with a resulting loss of consonant intelligibility. The effects of filtering are exacerbated by other distortions, particularly by background noise and hearing impairment.

Noise is the most common culprit, and its effectiveness as a masker depends on spectral, level, and temporal considerations. One of the most efficient maskers of speech is speech itself. To quote George Miller, "...the best place to hide a leaf is in the forest, and presumably the best place to hide a voice is among other voices." (Miller, 1947, p. 118). For this reason, a babble of many voices is often used in speech masking experiments. Broadband noise can also be an effective masker. At low-to-moderate levels of noise and speech, high-frequency noise masks more efficiently because it masks the consonant sounds, which are generally higher in frequency and lower in speech power than vowels. Figure 4, from Richards (1973), displays the relative energy of speech sounds. Because of a phenomenon known as the "spread of masking", low-frequency sounds become more efficient maskers as their intensity increases. Above a sound pressure level of approximately 80 dB, low-frequency masking increases at a faster than normal rate (Kryter, 1962a) and becomes increasingly effective at masking mid- and high-frequency sounds. Low frequency sounds, if intense enough, will mask the whole range of speech frequencies (Miller, 1947).

Noise becomes less efficient at masking speech when its levels vary with time. In its report on the effects of time-varying noise, CHABA (1981) points out that varying noise produces less speech masking than continuous noise for a given AI. The report predicts 97% sentence intelligibility for a time-varying L_{eq} of 70 dB, and even 81% intelligibility at an L_{eq} of 80 dB. The report also suggests that a good "speech interference index should be some running estimate that combines background noise and time-varying noise episodes in which the noise level is within 10 dB of the peak level." (CHABA, 1981, p. 7).

More often than not, distortions occur in combinations rather than singly. A talker may be smoking, chewing, or talking rapidly (Lacroix et al., 1979). He may have his head turned away, he may be trying to communicate at a distance, his vocal effort may be above or below the point of maximum intelligibility, or his articulation may be unclear. A very common combination of distortions is noise and low-pass filtering, which characterizes inefficient communication systems. Lacroix et al. (1979) investigated the effects of three types of distortion: increased rate of talking, interruption, and speech-shaped noise, singly, and in combination with low-pass filtering. The authors found that the reduction in speech recognition resulting from multiple distortions was considerably greater than

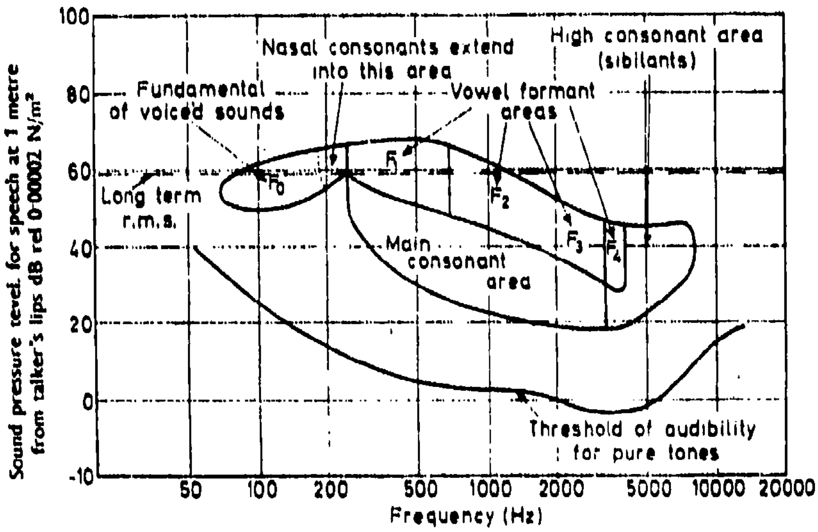

Figure 4. Relative energy of speech sounds.

Note. From <u>Telecommunication by Speech: The Transmission Performance of Telephone Networks</u> by D. L. Richards, 1973, London: Butterworths. Reprinted by permission.

an additive effect. According to Lacroix and his colleagues, these results corroborated similar findings of earlier investigations (Licklider and Pollack, 1948; Martin, Murphy, and Meyer, 1956; Harris, 1960).

III. TRANSMISSION CHARACTERISTICS

A. Distance Between Talker and Listener

Early criteria developed by Beranek (1950) gave estimated "speech interference levels" (SILs) as a function of distance and vocal effort. Based on communication in the free field, they show the expected 6-dB decrease in SIL for a given intelligibility with every doubling of distance. However, speech intelligibility will not deteriorate with distance as quickly as might be expected from the 6-dB per doubling rule because people will increase their vocal effort with increasing distance. Also, the 6-dB rule is inappropriate for indoor spaces because of room reverberation and other factors. Schultz (1984) has developed a formula for predicting sound propagation indoors, based on the frequency and sound power level of the source, room volume, and the distance from the source. Modifications to the SIL for vocal effort, reverberation, and other factors will be discussed in greater detail in a subsequent section.

Garinther and Hodge (1987) point out that individuals use a "communicating" voice level, meaning that they raise their voices as they feel necessary according to the distance at which they need to communicate. They cite research by Gardner (1966) to support their estimate of a 2.4 dB increase in vocal effort for each doubling of distance. An investigation of the effects of wearing a gas mask and hood showed that individuals use slightly higher voice levels in this condition, and raise their voices approximately 1.5 dB per doubling of distance (Garinther and Hodge, 1987).

B. Reverberation

Although reverberation is a necessary feature in concert halls and auditoriums, the prevailing thinking on the subject nowadays is that its effects on speech are virtually never beneficial. Early reverberations seem to have little adverse effect if they arrive during the production of the same sound (Nabelek, 1980), but Webster (1983) and other investigators he cites (Mankovsky, 1971; Kuttruff, 1973) believe that all reflections are detrimental. In a study of the influence of noise and reverberation on speech recognition, Nabelek and Pickett (1974) found that a change in reverberation time of 0.3 second produced a substantial decrease in speech recognition, equivalent to a 2- to 6-dB increase in noise level. The investigators used two types of noise: one consisting of 16 impulses/second and the other a babble of 8 talkers. Nabelek has reported that a degradation of speech perception in quiet occurs at reverberation times longer than 0.8 second, and that the amount of the degradation depends on the size of the room (and therefore the temporal distribution of reflections), the type of speech and noise, and the listener's distance from the source (Nabelek, 1980).

In an attempt to test the effects of small room reverberation and binaural hearing on normal and hearing-impaired subjects, Nabelek and Robinette (1978) found a significant decrease in speech recognition scores between a reverberation time of 0.25 to 0.5 second, and concluded that the adverse effects of reverberation are greater in small than they are in large rooms. A table comparing their data to those of other researchers shows that the effect of reverberation on speech recognition may vary anywhere from 0% to 34.8%, depending on reverberation time, presence or absence of noise, and monaural or binaural listening (Nabelek and Robinette, 1978, p. 246). The authors also discuss an experiment using computer simulated reverberation consisting of a direct sound followed by 5 reflections, decreasing at a rate of 6 dB per reflection. Unexpectedly, the results failed to show a statistically significant difference between speech recognition scores for three simulated reverberation times. In a later simulation, Nabelek (1980) did find a difference between non-reverberant and computer simulated reverberant conditions of 9% in the scores of hearing-impaired subjects. This simulation had been developed by Allen and Berkley (1979), whose FORTRAN program may be used to simulate a wide range of small-room acoustical conditions.

C. Spatial Location

The location of the speech and noise sources may also have an effect on speech intelligibility. The most difficult condition occurs when speech and noise are coming from the same direction. Generally, as the angle of separation becomes wider, intelligibility increases for a given speech-to-noise ratio. Plomp (1976) reports that with the speech signal coming from 0° azimuth, people could tolerate a decrease of approximately 5 dB in speech-to-noise ratio for the same intelligibility when the noise was moved from 0° to 135° azimuth. This finding occurred in non-reverberant conditions. Effects were less dramatic as reverberation time increased from 0 to 2.3 sec.

D. Monaural vs. Binaural Listening

Nature has provided us with two ears for reasons in addition to redundancy. Binaural hearing enhances our sense of a sound's location, and it increases our ability to recognize speech sounds in a reverberant space. We are able to do this by discriminating small differences in signal phase and time of arrival at the two ears. This ability is considerably better for frequencies below rather than above 1500 Hz (Littler, 1965).

Different investigators report different amounts of improvement or "binaural gain", defined as the difference in speech-to-noise ratio for a given speech recognition score. The amount of improvement depends on such aspects as reverberation, the type of masker, the spatial location of the speech and noise, the listener's hearing sensitivity, and the presence or absence of amplification. MacKeith and Coles (1971) report a 3- to 6-dB improvement from binaural summation alone (at or slightly above threshold). Nabelek and Pickett (1974) found improvements of 4 to 5 dB, unaided listening in reverberant conditions, but the gain was only 3 dB when listening through amplification. The binaural advantage appears to be greater for normal hearing than for hearing-impaired people (Nabelek and Robinette, 1978),

although the hearing-impaired will experience a peculiar summation when the hearing threshold levels for the two ears are dissimilar according to frequency (MacKeith and Coles, 1971).

Levitt and Rabiner (1967) have developed a method for predicting the gain in intelligibility due to binaural listening. They estimate the maximum benefit for single words in high-level white noise is about 13 dB, while at high intelligibility levels the benefit will be only about 3 dB (from summation). The authors suggest that binaural gain might be greater with speech as a masker, since Pollack and Pickett (1958) found advantages up to 12 dB. With respect to directionality, Plomp (1976) found that there was a binaural gain of about 2.5 dB over the monaural condition when the noise was on the side of the occluded ear, and a greater gain when the masking noise was on the side of the open ear. These advantages were fairly constant, irrespective of reverberation and azimuth of the masker. However, the data of Nabelek and Robinette (1978) and Nabelek and Pickett (1974) show sizeable increases in binaural advantage with a doubling of reverberation time.

E. Telephone Listening

Telephone circuitry filters the speech signal on both the low and high ends of the spectrum, such that the spectrum rises gradually from 200 Hz to a peak of about 800 Hz, with a gradual decline to 3000 Hz and a precipitous drop thereafter (Richards, 1973). Without the advantage of high-frequency speech information or binaural hearing, noise, either in the system or in the listener's environment, can be problematical. Noise in the listener's environment further disrupts telephone listening in that it is amplified through the same mechanism that enables talkers to monitor their voice levels, "side-tone feedback" (Holmes et al., 1983).

In an effort to evaluate the influence of a noisy background on telephone listening, Holmes et al. (1983) tested the ability of normal hearing subjects to hear speech through a standard "500" handset. Speech was presented at a sound pressure level of 86 dB (the average level of telephone speech according to the authors) in backgrounds of multi-talker babble and white noise at 65, 75, and 85 dB in five telephone conditions: transmitter off, transmitter occluded by the listener's palm, contralateral ear occluded, control (normal listening mode), and transmitter off plus contralateral ear occluded. The results showed no significant differences among conditions when the noise was at the 65 dB level, but for the less favorable speech-to-noise ratios, significantly poorer speech recognition scores were obtained during the control and contralateral ear occluded positions than during the transmitter off and transmitter occluded positions. The authors conclude that telephone listening can be improved by occluding the transmitter, but no help is derived from the popular remedy of occluding the opposite ear. Holmes and her colleagues also found that amplified telephones improve speech recognition because increases in the level of side-tone feedback are non-linear with respect to increases in signal level. They found that if the telephone's output was increased by as much as 20 dB, the side-tone feedback increased by only about 4 to 7 dB. Thus, the speech-to-noise ratio would be more favorable, and indeed they found that speech recognition scores using an amplifier showed smaller differences between the transmitter occluded and

unoccluded positions, causing the authors to recommend amplifier handsets as another remedy for telephone listening in noise.

F. Communication Systems

Communication systems have been specially designed for military and industrial use where high levels of background noise are common. Certain features have been developed to enhance the communication process in noise environments. Circumaural earcups house the receiver, providing attenuation of up to 20 to 30 dB, depending on frequency and on the effectiveness with which they are worn. The process of electronic peak clipping aids intelligibility by boosting consonant energy in relation to vowels, but the benefits of this process are limited when noise accompanies the signal (Kryter, 1984). The noise cancelling microphone is a useful innovation, as are improvements in circuitry such as the "expander/compander" circuitry described by Mayer and Lindburg (1981).

Despite recent improvements, Mayer and Lindburg (1981) contend that most communication systems in use today are based on design concepts that are over 50 years old. The 300-3000 Hz bandwidth allows insufficient intelligibility in noise, such that aviators sometimes need to take the time to use the phonetic alphabet - time that they can ill afford to spend. Mayer and Lindburg state further that peak clipping in typical noisy conditions can produce a distortion of the signal of up to 50%, degrading speech intelligibility to the extent that all the gains from this process are lost. They cite a worst case condition where peak clipping can almost destroy the intelligibility of a high amplitude "panic message". In addition, they maintain that current test procedures are outmoded. The 6cc coupler is not appropriate for circumaural earcups. ASA standard 1-1975 procedures are inappropriate because real-world, high noise environments lead to a "pumping" action on the earcup, causing the ear cushion to be lifted off the ear, with resulting acoustical leaks. Finally, Mayer and Lindburg state that the equipment used to test the noise cancelling microphone (the Kruff Box) is not an adequate simulator of the aircraft noise environment.

Mayer and Lindburg (1981) proceed to describe their newly developed test procedures and communication system. The test consists of "real head" attenuation in pink noise with two microphones, one outside and one beneath the earmuff. The system, C-10414 ARC Intercommunication Control has an increased bandwidth (300-4500 Hz) with a relatively flat response, and uses "expander/compander" circuitry, fast-acting automatic gain control, and a noise cancelling microphone. This kind of research and development will be continued under a program entitled The Voice Recognition and Response for Army Aircraft (VRAA).

Such a program also exists in the Air Force. The Voice Communication Research and Evaluation System (VOCRES) has been described by McKinley (1981) as a laboratory system replicating cockpit communication and environmental conditions, where the elements that can be varied include: microphones, earphones, helmets, oxygen masks, aircraft radios, ambient noise, jamming signal type and modulation, jammer-to-signal power ratios, and receiver input data.

IV. TALKER AND LISTENER VARIABLES

A. Talker Variables

1. Vocal effort and fatigue

Although people readily raise their voices in a noisy background or when separated by distance, there is a limit to the length of time they can and will maintain an increased vocal effort. Pickett (1956) identified the highest level, measured at one meter, that could be sustained without painful voice fatigue as 90 dB, and, regardless of fatigue, the highest absolute level was 100 to 105 dB. However, as Webster and Klumpp (1962) have indicated, people will be reluctant to expend a vocal effort beyond 78 dB for more than a brief period of time, even in higher noise levels. They call this the "asymptotic speech level" (Webster and Klumpp, 1962).

Rupf (1977) assessed subjective estimates of the length of time people could talk in noise before their voices would become unduly strained. He found that on the basis of 5-minute conversations in noise, about half the people believed they could talk for one hour in A-weighted levels of 75 dB, 30 minutes in 80 dB, 15 minutes in 85 dB, and 7 minutes in 90 dB. However, when asked to rate the feasibility of conversing during these 5-minute segments, the 50% level of acceptability fell at an A-weighted level of 83 dB.

Discomfort is not the only adverse effect of talking in high noise levels. Reports of noise-exposed workers show an abnormally high incidence of vocal cord dysfunction (vocal nodules, chronic hoarseness, etc.) among workers who need to communicate as part of their work (Anon., 1979; Klingholz et al., 1978; Schleier, 1977). Klingholz et al. (1978) found that 70% of laboratory subjects produced "pathological phonation" in A-weighted noise levels of 90 dB and above, and virtually all subjects did in levels above 95 dB. Clinical evidence of vocal disorders in noise-exposed workers with speech-intensive jobs showed that the disorders tended to occur between the third and seventh year of work (Klingholz et al., 1978).

2. Talker articulation

The talker's speech patterns can have considerable influence on the intelligibility of speech. Common sense tells us that people with a foreign accent, strong regional dialect, or just generally sloppy articulation will be more difficult to understand than people with standard dialect and careful enunciation. Borchgrevink (1981) alludes to potential air traffic safety hazards when controllers speak in foreign accents, "with errors in phoneme pronunciation and prosodic features." (p. 15-3). Picheny et al. (1985) gave a short review of the benefits gained by training personnel to articulate clearly. They cite Snidecor et al. (1944) as finding that drilling subjects to mimic the speech of a trained talker, as well as prompting them to talk louder, more clearly, and to open their mouths more, improved communication over military equipment. Similarly, Tolhurst (1955) was able to improve the intelligibility of speech in a noisy background by 10% when the talkers were instructed to speak more intelligibly. In another experiment, Tolhurst (1957)

found that by either decreasing speech rate or by increasing clarity, he was able to improve intelligibility by as much as 9% (see Picheny et al., 1985).

Picheny and his colleagues (1985) studied the effects on hearing-impaired listeners of conversational versus clear speech. Listeners were presented via headphones with short nonsense sentences at comfortable listening levels. When using the clear speech mode, talkers were instructed to enunciate consonants carefully, to avoid slurring words together, and to place stress on adjectives, nouns, and verbs. They were encouraged to talk as if they were speaking to a hearing-impaired listener in a noisy environment. Although listeners reported that the clear speech was tiring because it was spoken more slowly (sentences were approximately twice as long in the clear speech mode), the average improvement in intelligibility scores was 17%.

In a second article on the subject of clear speech, Picheny et al. (1986) presented an acoustical analysis of clear speech and the differences between the clear and conversational speech modes. They found that the increase in clear speech duration is achieved by lengthening the individual speech sounds as well as by inserting or lengthening pauses. They also found that clear speech is characterized by the consistent articulation of stop-burst consonants, and all consonant sounds at the end of words, both voiced and unvoiced. Although changes in the long term speech spectrum were small, the intensity for obstruent sounds (breath obstructed), appears to be up to 10 dB greater in clear than in conversational speech. The authors note that to date there is no hard evidence that isolates the most important acoustical factors in differentiating between clear and conversational speech. Consequently, they suggest the development of a model that will permit the synthetic manipulation of variables known to be important. In this way, "one could gradually transform conversational speech into clear speech by varying one parameter at a time...." (Picheny et al., 1986, p. 444).

Mosko (1981) studied the effect of clear speech on radio voice communications with normal listeners. Listeners were trained to "over-articulate" for a period of 3 to 4 days. For speech material, Mosko chose digit sequences and words that commonly occur in aircraft communications, presented in quiet and in noise. Again, the duration of the clear speech segments was up to twice as long as the normal utterances, and intelligibility showed a 16% to 18% improvement in quiet. Preliminary data from the noise conditions showed an improvement of 6% to 8% at a speech-to-noise ratio of 0 dB. In the discussion following his paper, Mosko points out that the speech of people using radio communication systems tends to deteriorate over time.

> You can almost chart how long they have been on the job by the deterioration in their speech and you notice this time and time again. When you train people to use radios...they should be professional talkers. (Mosko, 1981, p. 4-6)

3. Gender

There has been some controversy about the relative intelligibility of male and female voices. While the female voice is probably no less intelligible in most circumstances, it may be somewhat more difficult to

understand in high noise levels when it is lower in sound energy. Pearsons et al. (1977) found the female voice to be 2 dB lower than the male voice in the "casual," "normal," and "raised" modes, 5 dB lower in the "loud" mode, and 7 dB lower in "shout". They maintained that their data did not support Beranek's (1954) recommendation that background noise be reduced consistently by 5 dB to accommodate female talkers. In a study of speech materials processed through Air Force communication systems, Moore et al. (1981) found small but systematic differences in the intelligibility of male and female voices in high levels of background noise. While there was little difference at sound pressure levels of 79 and 95 dB, male voice intelligibility was 6.8% greater in 105 dB and 9.5% greater at a noise level of 115 dB. The authors were not sure whether the cause was that the high-frequency content of female speech was more easily masked, or because of the differences of vocal output with increasing levels of background noise.

B. Listener Variables

1. Preferred listening levels

Although quite high levels of speech can be tolerated with little or no loss of intelligibility if the speech is amplified and if the speech-to-noise ratios are sufficiently high, people prefer to listen to speech within a certain range of levels. A study by van Heusden et al. (1979) explores the relationships between selected listening levels in the sound field for speech and background noise. Using a Bekesy "up-down" adjustment method, listeners were instructed first to find the preferred speech level, as if listening to a radio, and later to find the minimum required level for understanding speech. (No details are given for the criteria for "understanding".) Speech and speech-shaped noise were presented through separate loudspeakers. A-weighted noise levels were 40, 50, 60, and 70 dB and quiet. The results showed average preferred speech levels of 49 dB(A) in quiet, and 61 dB(A) in noise, with a slope of 3.1 dB per 10 dB increase in background noise level --above about 35 dB(A). "Minimum" speech levels were identified as 25 dB(A) in quiet, and about 54 dB(A) in the 70-dB(A) noise condition, with a slope of 6.4 dB per 10 dB increase in noise level above 40 dB(A). The investigators concluded that people prefer to keep about the same (subjective) loudness level of speech in noise as they experienced in quiet, although this level will not guarantee the same level of intelligibility.

In a follow-up study by Pols et al. (1980), the same group of experimenters studied preferred listening levels for speech as a function of modulation frequency in fluctuating noise. Experimental conditions were similar, except that the noise, which was typical of community noise, was modulated at frequencies of 0.1, 0.3, 1 and 5 Hz. Also, subjects used a slightly different psychophysical method, which gave them somewhat more time in which to make their selections. The results showed that modulation frequency had a negligible effect on the selection of preferred listening level, so long as the equivalent sound level was constant among noise stimuli. However, the identified preferred levels were about 10 dB higher than in the previous experiment, and the slope of the curve was 5 dB per 10 dB increase in noise level above 35 dB(A), rather than 3.1 dB. Pols and his colleagues offer no explanation for the difference in slope, but they believe that the difference in level may be due to the difference in adjustment methods. In

this experiment, the method may have led to the identification of the most comfortable listening level, whereas in the previous experiment the levels identified would have reflected the just comfortable level. Pols et al. hypothesize a similar explanation for other such discrepancies they noted in the literature. This leads them to conclude that preferred listening levels are better described by a range of levels than by single numbers.

Investigations of preferred listening levels under earphones have produced somewhat higher levels, but there is considerable variation among studies. Beattie et al. (1982) measured most comfortable listening levels (MCL) in quiet, and in white noise levels of 55, 70, 85, and 100 dB SPL. The slope of the MCL, 5.3 dB per 10-dB increase in noise level, was similar to that of Pols et al. (1980), but the mean identified levels were much higher: 82.5 dB in quiet, and 90.9 dB and 100.3 dB in noise levels of 85 dB and 100 dB respectively. Beattie and his coworkers point out that there is a wide range of MCLs reported in the literature, varying from a low of 42 dB SPL in a study by Schaenman (1965) to a high of 91 dB found by Loftiss (1964). The discrepancies seem to be due mainly to differences in instructions and psychophysical methods of threshold determination (Beattie et al., 1982). One factor that would account for a portion (about 6 dB) of the difference between the results of Beattie et al. and the work of van Heusden et al. (1979) and Pols et al. (1980), is the difference in thresholds of sensitivity between listening in the sound field and under earphones. Another factor would be the use of A weighting by van Heusden and Pols, which would account for an additional 4 dB when compared with unweighted sound pressure levels, and still another is the use of higher noise levels by Beattie et al., which would be likely to induce listeners to raise speech levels.

In the above experiments, subjects were presented with a fixed level of noise and were permitted to adjust the preferred listening level separately. In a subsequent experiment (Beattie and Himes, 1984), subjects were presented with a fixed speech-to-noise ratio (under earphones) and asked to identify MCLs, adjusting the speech and noise together, as they would when listening through a hearing aid or a communication system. The investigators found MCLs that ranged from 78 dB SPL in a speech-to-noise ratio of -10 to 83 dB SPL in a speech-to-noise ratio of +10. Upper ranges of comfort, defined as the point at which listening would be uncomfortable if the level were any louder, were 93 dB SPL for a speech-to-noise ratio of -10, and 98 dB SPL in quiet. Although there was a great deal of individual variability, it is interesting to note that people will include higher levels of speech within the comfort zone, so long as they do not have to contend with too much noise.

2. Non-native listeners

Degraded communication can occur when listeners, as well as talkers, use a language which is not their native tongue. Using short, high-predictability and low-predictability sentences (the SPIN test), Florentine (1985) tested 11 native and 14 non-native but fluent-in-English listeners. She found that the native listeners were able to obtain 50% performance levels at significantly lower speech-to-noise ratios (about 3 dB) than the non-native listeners. Likewise, Nabelek (1983) found differences between native and non-native listeners as a function of reverberation. In a reverberation time of 0.4 second, non-natives scored 6% lower, and with reverberation times of 0.8 to 1.2 second, they scored 10% lower than native listeners.

In an interesting study of the effects of listening in a second language, Borchgrevink (1981) selected 13 Norwegian men who were fluent in English and 13 Englishmen fluent in Norwegian, and tested them with "everyday" Norwegian and English sentences, balanced and matched for syntax and phoneme frequency. The subjects listened to these sentences at 65 dB SPL in the sound field, in a noise background that was decreased between sentence sets in 2-dB steps from 76 to 56 dB SPL. The results showed that both the Norwegian and English subjects needed significantly more favorable speech-to-noise ratios when listening in their second language (see Table 1).

Table 1

Speech-to-Noise Ratios Needed for Correct Repetition of Sentences
(From Borchgrevink, 1981)

	S/N Needed by Norwegian Subjects		S/N Needed by English Subjects	
	Mean	SD	Mean	SD
Norwegian Sentences	0.1	1.28	3.9	2.66
English Sentences	2.1	1.66	0.4	1.36

The author notes that individuals need fewer acoustical cues to understand sentences presented in their first language, even when they are fluent in the second language. He concludes that subjects are better equipped to synthesize a degraded message in their native language because of a more firmly established "concept-reference coherence" (Borchgrevink, 1981).

3. Speech recognition during a secondary task

Although the literature on the subject is not extensive, it appears that noise has an added disruptive effect when the listener must comprehend speech and perform another task simultaneously. Lazarus (1983) presents data from Hormann and Ortscheid (1981) showing that speech recognition scores decrease as a function of speech-to-noise ratio more rapidly when a visual memory task is added. Jones and Broadbent (1979) cite other investigations indicating that subjects trying to understand speech in noise have difficulty remembering material learned in quiet (Rabbitt, 1966, 1968). They describe an earlier experiment by Broadbent (1958) in which subjects were presented with speech in noise, and with speech filtered as if it were masked by noise. While there was no significant difference in speech recognition scores, there was a deficit in a secondary tracking task in the noise condition which did not occur in the filtered speech condition. The authors conclude that the

extended effort required to cope with the noise produces a penalty in other activities (Jones and Broadbent, 1979).

4. Auditory fatigue

In this context, auditory fatigue may mean temporary threshold shift (TTS) or a more central effect "analogous to perstimulatory fatigue or loudness adaptation" (Pollack, 1958). Regardless of the etiology, high noise or speech levels may produce a deterioration in speech recognition with continued exposure.

Pollack (1958) investigated the effects of broadband noise and speech levels (S/N = 0 dB) of 110 dB to 130 dB for successive 100-second exposures. Speech recognition scores deteriorated significantly over successive tests at noise and speech levels above 115 dB, and the deterioration in time was roughly logarithmic over the period of the eight tests. Not unexpectedly, post-stimulatory tests showed large decrements in speech recognition for soft (45 dB) and very loud (125 dB) speech, but no significant effects on speech in quiet between these levels (Pollack, 1958).

In another study of the effects of auditory fatigue, Parker _et al._ (1980) exposed subjects to a 1500- to 3000-Hz band of noise at 115 dB for 5 minutes. After noise exposure, recognition scores for PBs in a 2825- to 3185-Hz band of noise were poorer in quiet, slightly poorer in the 90-dB noise condition, about the same in 40 dB, and somewhat better in the 65 dB noise condition. The authors conclude that the subjects responded as predicted from a "recruitment model" (referring to the improvement in the 65 dB noise condition, and suggest that a small TTS would not affect speech embedded in moderately intense masking noise.

Sorin and Thouin-Daniel (1983) studied the effects of mild TTS on the recognition of low-level speech in noise (speech at 34 dB(A), noise at 40 dB(A)). They added a "lexical decision" task in the form of a word/non-word judgement, in an attempt to test central as well as peripheral dysfunction. The results showed that the presence of a 15-dB TTS produced an increase from 5.3% to 10.8% incorrect rhyme words and 5% to 13.5% incorrect lexical responses (which includes decisions exceeding a 2-second limit). They also noticed that the presence of TTS increased a subject's tendency to respond "word" more often than "non-word", a type of response that has been identified in studies of the effects of noise on task performance. Although speech at 34 dB(A) is not typical of everyday conversation, it could characterize certain combat conditions, where understanding softly spoken messages is of vital strategic importance.

V. PREDICTION METHODS

A. Articulation Index

The Articulation Index (AI) is a method for predicting the efficacy of speech communication in noise, based on the research and method of French and Steinberg (1947). The classic "20 band" method uses measurements or estimates

of the spectrum level of speech and noise in 20 contiguous bands, each of which contribute equally to speech intelligibility. This method has been improved and modified by Kryter and his colleagues for numerous conditions of noise and distortion (see Kryter, 1962a and ANSI, 1969). These modifications include:

1. Corrections for reverberation times up to 9 sec.

2. Corrections to the noise spectrum for spread of masking effects (upward, downward and nonlinear growth).

3. Methods using octave and 1/3 octave bands instead of the original 20 bands.

4. Calculation of AI for non-steady-state noise with a known duty-cycle and levels that fall at least 20 dB during the "off period".

5. Calculation of AI for non-steady noise when the rate of interruption is known.

6. Adjustments for the effects of sharp, symmetrical peak clipping.

7. Corrections for vocal effort, including speech levels of 40 to 100 dB (long-term rms).

8. Corrections for the added benefits of lipreading.

Applications of the AI to hearing-impaired listeners have been suggested by Kryter (1970), Braida <u>et al.</u> (1979), Dugal <u>et al.</u> (1980), Skinner and Miller (1983), Kamm <u>et al.</u> (1985), and Humes <u>et al.</u>, (1986).

Although the AI can be somewhat complicated in terms of measurement and instrumentation, it has been found to be a valid predictor of speech intelligibility in a variety of conditions (Kryter, 1962b), and it has been a popular and a respected measurement tool over recent decades.

B. Speech Interference Level

Originally developed by Beranek (1954), the Speech Interference Level (SIL) provides a quick method of estimating the distance with which communication can occur for various levels of vocal effort. The current method involves taking the arithmetic average of sound levels in the octave bands 500, 1000, 2000, and 4000 Hz. According to ANSI S3.14 (ASA, 1977), the primary purpose of the SIL is to rank-order noises with respect to speech interference. Figure 5, from ASA (1977), shows talker-to-listener distances for "just reliable" communication (defined as 70% monosyllables), with the approximate A-weighted level on the abcissa for comparison. "Expected voice level" reflects the natural increase in vocal effort with increasing SIL.

Figure 6 shows Webster's most recent version of the SIL criteria (Webster, 1983), with numerous modifications and embellishments. Webster (1983), describes them as: (1) a broader range of voice levels to reflect differences between public and private voice levels (see Houtgast, 1980; van

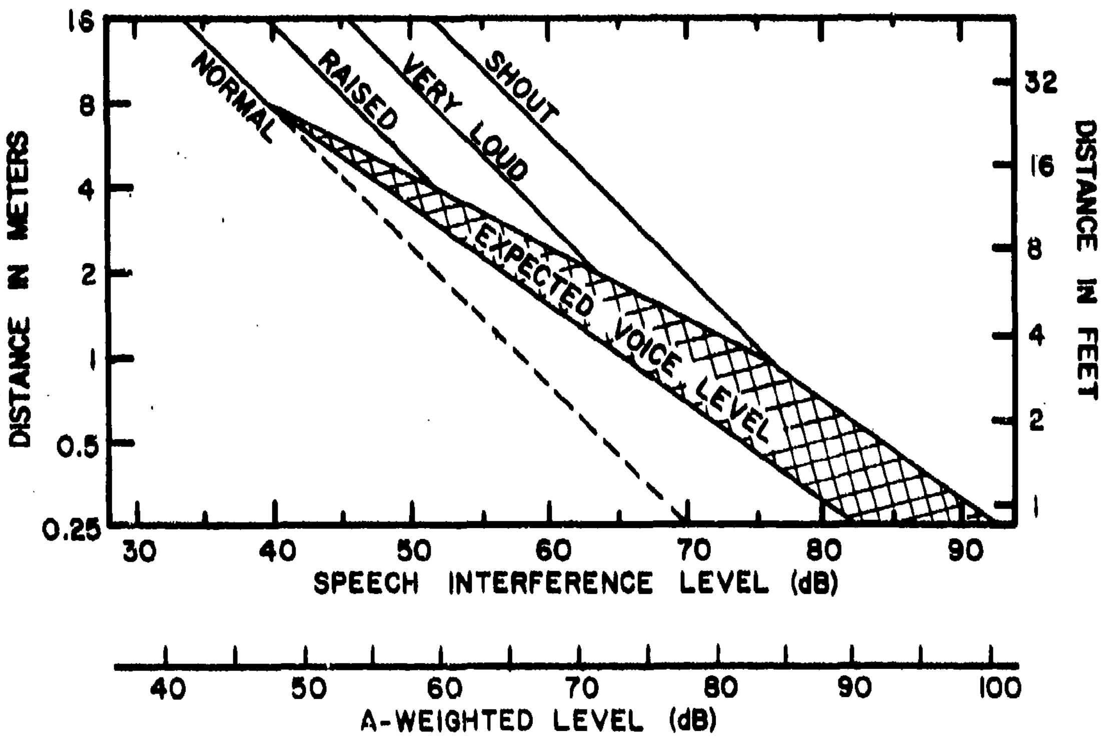

Figure 5. Talker-to-listener distances for just reliable communication.

Note. Excerpted from ANSI S3.14-1977 (Revised 1986) American National Standard for Rating Noise With Respect to Speech Interference, Acoustical Society of America, 335 East 45th Street, New York, NY 10017. Reprinted by permission of the Acoustical Society of America, New York, New York.

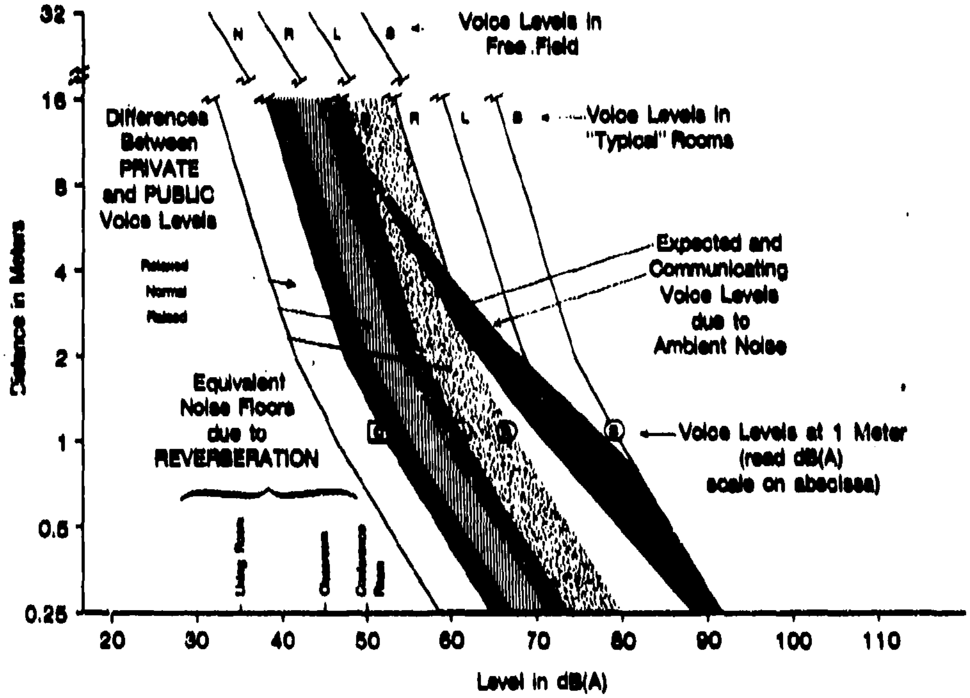

Figure 6. Revised "SIL" chart showing relationships among A-weighted ambient noise levels, distances between communicators, and voice levels of talkers for just reliable communication indoors.

Note. From "Communicating in Noise, 1978-1983" by J. C. Webster, in G. Rossi (Ed.), Noise as a Public Health Problem, 1983, Proceedings of the Fourth International Congress, Milan, Italy: Centro Ricerche e Studi Amplifon.

Heusden et al., 1979); (2) a different rate of fall-off of speech level with distance based on typical room reverberation; (3) "equivalent noise floors" based on room reverberation (see Houtgast, 1980); and (4) a downward shift of 3 dB in the voice level reference lines at one meter to account for the differences between A-weighted and rms speech levels (according to Steeneken and Houtgast 1978). Despite all of these modifications, the SIL still has certain disadvantages in that it assumes normal hearing on the part of the listener, and face-to-face communication with unexpected word material (Webster, 1984), and it uses only one level of intelligibility (70% monosyllables). Someone who desired 90% word intelligibility, for example, would not be able to use the chart.

C. Speech Transmission Index

Developed by a group of researchers at the TNO Institute for Perception in the Netherlands, the Speech Transmission Index (STI) is derived from a speech transmission channel's "Modulation Transfer Function" (MTF). The MTF may be measured with special equipment or calculated from the volume and reverberation time of the room, distance between talker and listener, and noise level (Houtgast, 1980). Figure 7, from Houtgast (1980) shows a model of the derivation of the STI. Houtgast gives data indicating an excellent correlation between STI and speech intelligibility for a wide variety of large rooms. The author also explains that in a highly reverberant room, noise below a certain level can have no degrading effect on speech because the adverse effects of reverberation dominate. This is the "noise floor", which Webster has incorporated in his latest SIL chart (see Figure 6).

In a later paper, Houtgast and Steeneken (1983) discuss the verification of the original model, which had used only speech-shaped noise, reverberation, and Dutch monosyllables. Subsequent research showed the STI to be a good predictor of speech intelligibility (1) in five types of noise spectra; (2) with other distortions besides reverberation, such as filtering, peak-clipping, and automatic gain control; (3) with untrained subjects outside the laboratory; (4) for sentences in addition to monosyllables; and (5) for seven other languages besides Dutch (Houtgast and Steeneken, 1983).

Humes et al. (1986) modified the STI by analyzing spectral information from the speech and noise signals in one-third octave rather than octave bands, and by weighting the bands according to the method originally developed by French and Steinberg (1947) for the AI. Humes and his colleagues found that these adjustments improved the STI's ability to predict speech recognition scores in both normal-hearing and hearing-impaired listeners.

In a subsequent effort, Humes et al. (1987) tested their modified STI (mSTI) on a large set of existing speech recognition data obtained under a variety of conditions, including low-pass and high-pass filtering, and various speech levels and speech-to-noise ratios. They found that the mSTI was a good predictor of speech recognition in all conditions, with the exception of low-pass filtering. The investigators speculate that increasing the frequency resolution of the mSTI (and AI) from 15 to 20 bands might solve this problem.

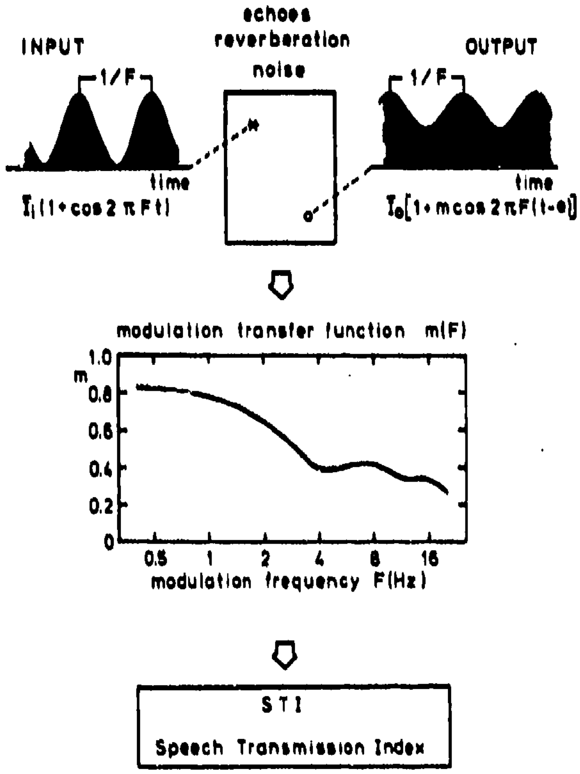

Figure 7. Derivation of the Speech Transmission Index.

<u>Note</u>. From "Indoor Speech Intelligibility and Indoor Noise Level Criteria" by T. Houtgast, in J. V. Tobias, G. Jansen, and W. D. Ward (Eds.), <u>Proceedings of the Third International Congress on Noise as a Public Health Problem</u> (ASHA Reports 10), 1980, Rockville, MD: American Speech-Language-Hearing Association. Reprinted by permission.

D. Sound Level Meter Weighting Networks

Aside from the fact that the sound level meter with its A-weighting network is inexpensive, readily available, and easy to use, it is a good predictor of speech interference, especially in noise spectra that are not unduly complex. Klumpp and Webster (1963) found A-weighting far superior to the other weighting networks, and Webster has effectively substituted A-weighting for SIL in his latest "SIL" chart (see Figure 6). Measuring the noise, however, gives only part of the information of interest. The A-weighting network can also be used effectively to predict AI and STI by measuring both speech and noise levels to obtain a speech-to-noise ratio. In addition, Webster (1984) also points out that A-weighting is amenable (as are all weighting networks) to time integration. Second to the AI, CHABA Working Group 83 recommends the A-weighted L_{eq} for predicting the effects on speech intelligibility of time-varying noise (CHABA, 1981).

Based on an analysis of 16 equally speech-interfering Navy noises (Klumpp and Webster, 1963), Webster (1964) developed a set of speech-interference (SI) contours which could serve as sound level meter weighting networks. In this process, Webster found that as the AI (and consequently speech intelligibility) increased, the frequencies that most effectively mask speech increase from about 800 Hz to around 3000 Hz (Webster, 1964). Figure 8 shows Webster's SI curves, including two curves originally developed by Beranek (1957). The S-I 50 curve is appropriate for an AI of 0.8, the S-I 60 for an AI of 0.5, S-I 70 for an AI of 0.2, and the S-I 80 for an AI of up to 0.05. Although these curves have never been incorporated into standard sound level meters, they would seem to offer some interesting possibilities.

E. Relationship of Methods to One Another

These predictive methods can be viewed together with respect to their physical interrelationships, and to their relative merit as predictors. ANSI S3.14 (ASA, 1977) states that for many common noises, the SIL (yielding 70% intelligibility) will be about 8 dB below the A-weighted sound level. According to ANSI S3.5 (ANSI, 1969), 70% monosyllable intelligibility (for 1000 PBs) is achieved at an AI of 0.45, which translates to an approximate speech-to-noise ratio of 1.5 dB. For speech-shaped noise, the STI and AI have a uniform and predictable relationship. A speech-to-noise ratio (S/N) of 1.5 dB corresponding to an AI of 0.45 will yield an STI of 0.55 (see Houtgast, 1980). This relationship can be seen as:

$$AI = (S/N)/30 + 0.4$$

$$STI = (S/N)/30 + 0.5$$

To assess the effectiveness of various rating schemes, Klumpp and Webster (1963) compared AI, dB(A), two versions of the SIL, and various other measures in 16 equally-interfering Navy noises. They found that the AI showed the least variability, followed by the SIL 355 Hz to 2800 Hz, dB(A), and SIL 600 Hz to 4800 Hz. Kryter and Williams (1965) found that the SIL 600 Hz to 4800 Hz outperformed the SIL 355 Hz to 2800 Hz in aircraft noises, which generally contain a greater proportion of high frequencies than the Navy noises.

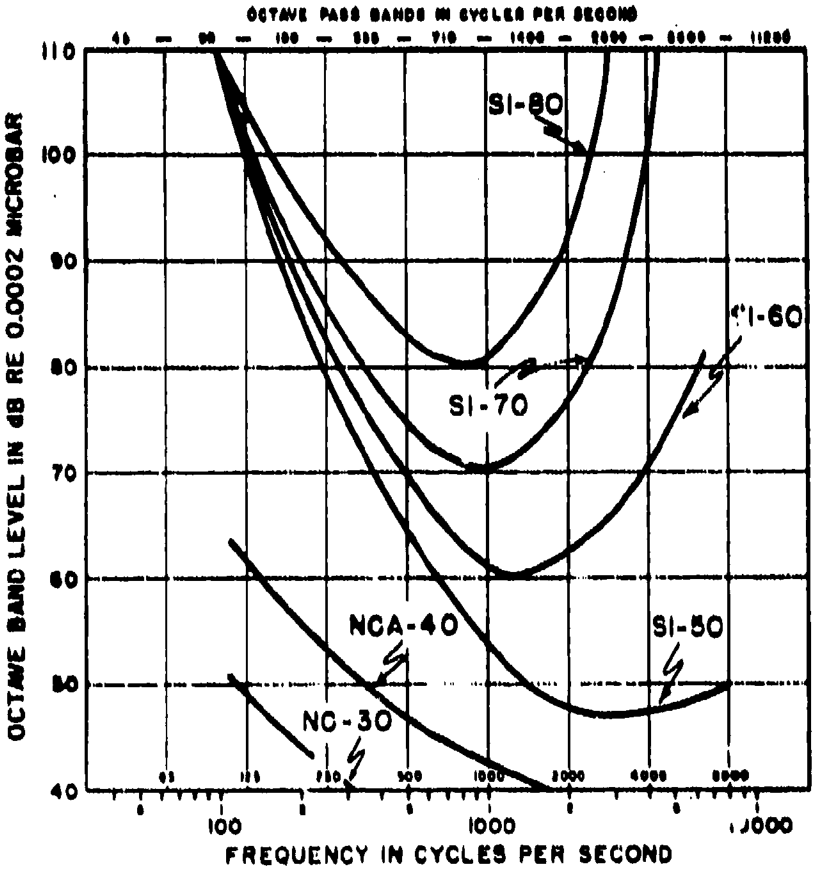

Figure 8. Webster's suggested speech interference (SI) contours, including two noise criteria curves from Beranek.

Note. From "Relations Between Speech-Interference Contours and Idealized Articulation-Index Contours" by J. C. Webster, 1964, Journal of Acoustical Society of America, 36, pp. 1662-1669.

In a recent study, Bradley (1986) compared four methods for predicting speech intelligibility in medium-sized to large rooms: AI, A-weighted speech-to-noise ratio, STI, and Lochner and Burger's (1964) "useful/detrimental" sound ratios. In this latter method, useful energy is defined as the weighted sum of energy arriving in the first 0.095 second after the arrival of direct sound. Detrimental energy is any later-arriving energy from the speech source plus background noise in the room (Bradley, 1986). The results showed that all methods did reasonably well, but the Lochner/Burger method produced the highest correlation with speech intelligibility and the lowest error. The AI and A-weighted speech-to-noise ratio performed nearly as well, and the STI ranked fourth in effectiveness. The author concludes that a satisfactory and simple approach would be to measure the A-weighted speech-to-noise ratio and the reverberation time at 1000 Hz, and use the regression coefficients he developed to form prediction equations (see Table I and Figure 9 in Bradley, 1986).

VI. ACCEPTABILITY CRITERIA

A. Minimal or "Just Reliable" Communication

There is a paucity of information on the subject of communication requirements for specific activities. Quite a few investigators refer to minimum requirements for "just reliable" communication, but few elaborate on the uses of this level of communication, or on the amount of communication needed for various purposes. Most agree that the minimum conditions to barely communicate range from an AI of 0.3 to 0.45. Table 2 gives recommendations for "just reliable" communication conditions from five sources. Data actually mentioned by the sources are underlined, and the remaining data have been filled in with the help of ANSI S3.5 (1969) (see Figure 3).

Although ANSI S3.5 gives no specifications for "just reliable" communication, the standard states:

> What level of performance is to be required over a given system is, of course, dependent upon factors whose importance can be evaluated only by the users of the communication system. Present-day commercial communication systems are usually designed for operation under conditions that provide AI's in excess of 0.5. For communication systems to be used under a variety of stress conditions and by a large number of different talkers and listeners having varying degrees of skill, an AI of 0.7 or higher appears appropriate.
> (ANSI, 1969)

B. Recommendations for Various Environments and Operations

Although the literature is virtually silent on the specific amount and type of communication needed for various operations, there exist numerous

Table 2

Conditions for "Just Reliable" Communication. Underlined data are those identified by sources. Other data estimated using Figure 3 (Figure 15 in ANSI, 1969).

Source	AI	S/N	STI	Monosyllable Intelligibility	Sentence Intelligibility	Author's Comment
Beranek (1947)	<u>0.3</u>	-3 dB	0.4	43%[1]	80%[2]	Unsatisfactory or marginally satisfactory.
ANSI S3.14 (ASA, 1977)	0.45	1.5 dB	0.55	<u>70%</u>[1]	95%[2]	Just reliable.
Houtgast (1980)	0.4	0 dB	<u>0.5</u>	62%[1]	93%[2]	Just reliable.
CHABA (1987)	0.35	<u>-1.5 dB</u>	<u>0.45</u>	54%[1]	<u>88%</u>[2]	Lower range of "fair".
Webster and Allen (1972)[3]	<u>0.35</u>	-1.5 dB	0.45	<u>80%</u>[4]	<u>95%</u>	Minimum acceptable for certain military communication equipment operating in "highly adverse" conditions.

[1]1000 PB words
[2]Unfamiliar sentences
[3]Cited in Webster, 1978
[4]Fairbanks Rhyme Words

recommendations for background sound levels that are appropriate for certain activities and spaces. For example, the German government recommends the following "rating level" (A-weighted Leq with corrections for impulses and tones): maximum of 55 dB for jobs that involve mental activity; maximum of 70 dB for simple and mechanized office activities; maximum of 85 dB for all other activities (Lazarus, 1983). According to the U.S. Environmental Protection Agency (EPA, 1974), A-weighted background noise levels of 45 dB will allow 100% intelligibility of relaxed conversation indoors, and 95% sentence intelligibility is achieved at a level of about 64 dB.

Table 3 shows recommendations from Beranek et al. (1971) for "preferred noise criterion" (PNC) curves and A-weighted background noise levels to achieve various levels of communication in various types of spaces. Levels of 66 to 80 dB are recommended for work spaces where communication is not required. For the others, the recommendations range from 56 to 66 dB for "just acceptable" speech and telephone communication in shops, garages, power-plant control rooms, etc., to 21 to 30 dB for excellent listening conditions in large auditoriums and concert halls.

The National Aeronautics and Space Administration (NASA) asked the National Academy of Sciences/National Research Council's Committee on Hearing, Bioacoustics, and Biomechanics (CHABA) to draft criteria for speech communication aboard the future NASA Space Station (CHABA, 1987). In their report, the authors direct A-weighted speech levels of 62 dB in the direct field, and 60 dB in the indirect field (greater than one meter), where most of the communication would take place. To obtain a minimum speech-to-noise ratio of 5 dB, the maximum noise level should be 55 dB(A). Assuming a one-second reverberation time, this translates to an STI of 0.45, an AI of 0.35, and sentence intelligibility of 88% (see Table 2). The authors mention a recommended range of STIs from 0.45 to 0.6, which would yield sentence intelligibility of up to 95% (CHABA, 1987). Presumably, however, for an STI of 0.6, either the reverberation time would have to be reduced or the speech-to-noise ratio should be considerably higher.

C. Consequences of Degraded Speech

Although the consequences of degraded speech can be extremely serious, most of the references in the literature are anecdotal or subjective. While these kinds of findings lack the power of objective, quantified research results, they are nevertheless compelling. For example, Williams and his coauthors state: "Field reports have indicated situations wherein troops emplaning from rotary-wing aircraft sometimes experience hearing threshold shifts of such severity that they are unable to make use of aural cues in detecting enemy movements." (Williams et al., 1970, p. 1)

At a recent conference entitled Aural Communication in Aviation, sponsored by the Advisory Group for Aerospace Research and Development (AGARD), some of the contributors alluded to the consequences of degraded speech. Mayer and Lindburg (1981) pointed out that future battles will be fought on or near the ground in a "nap of the earth" environment. This will increase the aviator's already heavy workload. The fatiguing effects of high noise and poor communication will have adverse effects on aviators' combat

Table 3

Recommended preferred noise criteria (PNC) and A-weighted levels for steady background noise in various indoor areas.

Type of space (and acoustical requirements)	PNC curve	Approximate L_A, dBA
Concert halls, opera houses, and recital halls (for listening to faint musical sounds)	10 to 20	21 to 30
Broadcast and recording studios (distant microphone pickup used)	10 to 20	21 to 30
Large auditoriums, large drama theaters, and churches (for excellent listening conditions)	Not to exceed 20	Not to exceed 30
Broadcast, television, and recording studios (close microphone pickup only)	Not to exceed 25	Not to exceed 34
Small auditoriums, small theaters, small churches, music rehearsal rooms, large meeting and conference rooms (for good listening), or executive offices and conference rooms for 50 people (no amplification)	Not to exceed 35	Not to exceed 42
Bedrooms, sleeping quarters, hospitals, residences, apartments, hotels, motels, etc. (for sleeping, resting, relaxing)	25 to 40	34 to 47
Private or semiprivate offices, small conference rooms, classrooms, libraries, etc. (for good listening conditions)	30 to 40	38 to 47
Living rooms and similar spaces in dwellings (for conversing or listening to radio and TV)	30 to 40	38 to 47
Large offices, reception areas, retail shops and stores, cafeterias, restaurants, etc. (for moderately good listening conditions)	35 to 45	42 to 52
Lobbies, laboratory work spaces, drafting and engineering rooms, general secretarial areas (for fair listening conditions)	40 to 50	47 to 56
Light maintenance shops, office and computer equipment rooms, kitchens, and laundries (for moderately fair listening conditions)	45 to 55	52 to 61
Shops, garages, power-plant control rooms, etc. (for just acceptable speech and telephone communication). Levels above PNC-60 are not recommended for any office or communication situation	50 to 60	56 to 66
For work spaces where speech or telephone communication is not required, but where there must be no risk of hearing damage	60 to 75	66 to 80

Note. From "Preferred Noise Criterion (PNC) Curves and Their Application to Rooms" by L. L. Beranek, W. E. Blazier, and J. J. Figwer, 1971, Journal of Acoustical Society of America, 50, pp. 1223-1228. Reprinted by permission.

effectiveness (Mayer and Lindburg, 1981). Conference Chairman Money referred to the "significance of failure or inadequacy of speech communication or audio warning systems in military operations..." and the consequent "cost in training and reduction of operational effectiveness." (Money, 1981, p. ix). In a discussion of clear speech later in the meeting, McKinley remarked:

> Your references to standard language has prompted me to make these remarks about something I have found in examining tapes of last messages from pilots during accidents. It is that usually, the message is a short unfamiliar language and in many cases, unintelligible. I think they could have been intelligible if the system had been designed correctly.

One study that simulates the consequences of communication failures in terms of error rates is the study of speech recognition using the M25 gas mask by Garinther and Hodge (1987). The authors cite the Defense Department's MIL-STD-1472C (DoD, 1981) defining "minimally acceptable" communication as a PB score of 43%, and "normally acceptable" communication as a PB score of 75%. Gas mask wearers were unable to achieve the 75% level at a distance of only one meter, and the 43%, minimal, level was achieved at a distance of 12.5 meters. Unmasked listeners could achieve this level at approximately 48 meters. Garinther and Hodge note that 12.5 meters is about one-half the distance at which platoon leaders would like to be able to communicate in field conditions. On the basis of their data, they estimate that using maximum vocal effort at a distance of 12.5 meters, individuals wearing gas masks would have an error rate of 3% with a small set of standard words, 7% with standard, previously known sentences, and 20% with non-standard sentences. One could expect an even higher error rate for non-standard words out of context. Garinther and Hodge also point out that maximum vocal effort can be sustained for only a short period of time.

Any system that allows less than 100% intelligibility assumes that some words will be lost or misunderstood. Systems that are designed for "just reliable" or "fair" communication depend for an extra margin of safety upon the normal redundancy of sentences, and especially upon the added redundancy provided by standard phraseologies, such as air traffic control language. These systems will function relatively effectively under normal conditions. However, normal conditions may be disrupted by any number of causes: an emergency requiring a non-standard word; a sudden decrease in speech-to-noise ratio; a momentary equipment failure; or a "panic" situation in which intelligibility is drastically reduced. The consequences of inadequate or misunderstood instructions in these situations can be dire indeed: in the extreme, loss of life and destruction of expensive equipment.

VII. DETECTION OF WARNING SIGNALS IN NOISE

Noise can mask warning sounds in the same way it masks speech. Theoretically, a warning sound will be audible if any frequency in the sound exceeds the critical ratio with respect to the surrounding band of noise. But

because the signal is detectable does not necessarily mean it will be effective. In the development of criteria for audible warning signals, Wilkins and Martin (1982) differentiate between detectability, demand on attention, and recognizability of signals. They point out that inattention may elevate the masked thresholds of warning signals (over the threshold of detectability), and that an even greater signal-to-noise ratio could be necessary when the signal is embedded among other meaningful, but irrelevant stimuli. These investigators cite a level of at least 15 dB above masked threshold as a widely accepted safety margin, and advocate a signal-to-noise ratio of at least 18 dB for 100% detectability especially if hearing protection is used (Wilkins and Martin, 1982).

Coleman et al. (1984) concur with the need for a 15-dB difference between signal level and masked threshold to produce "clear audibility" (p. 21). They maintain that as the signal approaches this level the listener will regain perceptual abilities and the ability to localize the direction of the signal source. But the 15-dB difference does not guarantee the signal's ability to claim the subject's attention.

The National Fire Prevention Association asked CHABA to develop a national fire alarm signal (Swets et al., 1975). The criteria were that the signal must be easily detected above background noise, different from other alarm signals, and adaptable to existing systems. The CHABA working group recommended a standard temporal profile, consisting of two short bursts and a long burst. Nominal on-segments should be between 0.4 and 0.6 second and off-segments between 0.3 and 0.6 second, with a rise and decay of 10 dB within 0.1 second. The on state should exceed the listener's 24-hour L_{eq} by 15 dB, and should exceed by 5 dB any maximum level for which the duration is greater than 30 seconds.

The working group cautioned users not to exceed a level of 30 dB without "consultation with local health authorities".

In a very thorough and well researched effort, Patterson (1982) offers a set of guidelines for auditory warning systems for civil aircraft. He has identified numerous problems with existing civil aviation warning systems:

1. The warning levels are too loud. "They flood the flight-deck with very loud, strident sounds," disrupting thought patterns and communication, and making the systems unpopular with the crews (p. 1).

2. Temporal characteristics are unsatisfactory. The onsets and offsets are sufficiently abrupt to evoke startle reactions, the temporal patterns are not sufficiently distinctive, and the total on-times are too long, interfering with speech communication.

3. Low priority warnings sometimes appear more urgent than high priority warnings.

4. The ergonomics of these warning systems are "deplorable". They are lacking in a sense of perspective, meaning that many are false and others have confused priorities. The aversive character of the sound is likely to convince the crew to cancel it as quickly as possible, thereby canceling the protection it provides.

5. Voice warnings are not frequently used, and the speech quality of existing systems is not good.

To correct these defects, Patterson developed a prototype warning system based on a comprehensive research effort. The following guidelines resulted:

1. Overall level should be at least 15 dB and not more than 25 dB above masked threshold.

2. The temporal pattern should consist of pulses with 20 to 30 msec rise and decay times, and gating functions that are rounded and concave downward. Pulse duration should be 100 to 150 msec, and intervals between pulses should be less than 150 msec for urgent and greater than 300 msec for non-urgent warnings. Each warning "burst" should consist of a set of 5 or more pulses in a distinctive temporal pattern.

3. The spectrum should consist of 4 harmonically related components between the frequencies of 500 to 5000 Hz, with a fundamental frequency between 150 and 1000 Hz. Signals demanding immediate action should contain a few quasi-harmonic components and/or a brief frequency glide.

4. For ergonomic reasons, manual volume control should be avoided, and AVC should be restricted to a 10 to 15 dB range. The total repertoire of signals should consist of not more than 6 immediate-action signals and up to 3 "attensons" (less urgent but attention demanding sounds like musical chords).

5. Voice warnings for immediate action should be brief, without repetition, and in a key-word format. Less urgent warnings can be full phrase and can be repeated. The system should accommodate a frequency range of 500 to 5000 Hz, and there should be progressive amplification of 3 dB/octave between those frequencies.

Figure 9 gives the component patterns for an "advanced" warning signal developed by Patterson (1982), showing four regularly spaced pulses followed by two irregularly spaced pulses. Sound level is reflected by the ordinate, and time is on the abcissa. Rows 3 and 4 show increasing levels of urgency. Figure 10, also from Patterson (1982), shows the time course of a complete warning. Each little trapezoid represents a series of pulses as in Figure 9, and the relative intensities are reflected by trapezoid height. This warning includes the voice message "undercarriage unsafe".

Subsequent to their development, Patterson's guidelines have been used with both conventional and rotary-wing aircraft, as well as in hospitals (Patterson, 1985). Rood et al. (1985) have adapted Patterson's guidelines to the conditions found in military helicopters. The authors point out certain differences between helicopters and civil aircraft. The pace of life on a helicopter "flight deck" is much faster than in a civil airliner, and the noise spectrum is different. Rood and his colleagues recommend a double burst of an attenson followed by a voice warning, with repeats as appropriate. Each primary warning has its own "attenson", made distinctive by pulse and burst parameters, with urgency controlled both by spectral and temporal characteristics. Spectral characteristics are matched to the particular aircraft, helmet, and the response characteristics of transducer in use (Rood et al., 1985).

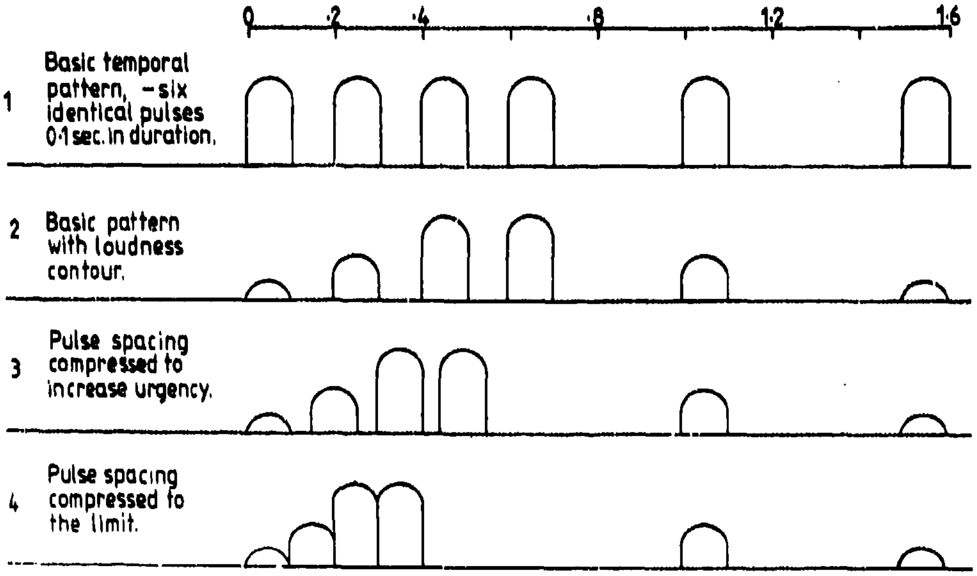

Figure 9. Component patterns for an advanced auditory warning signal.

Note. From Guidelines for Auditory Warning Systems on Civil Aircraft (CAA Paper 82017) by R. D. Patterson, 1982, London: Civil Aviation Authority. Reprinted by permission.

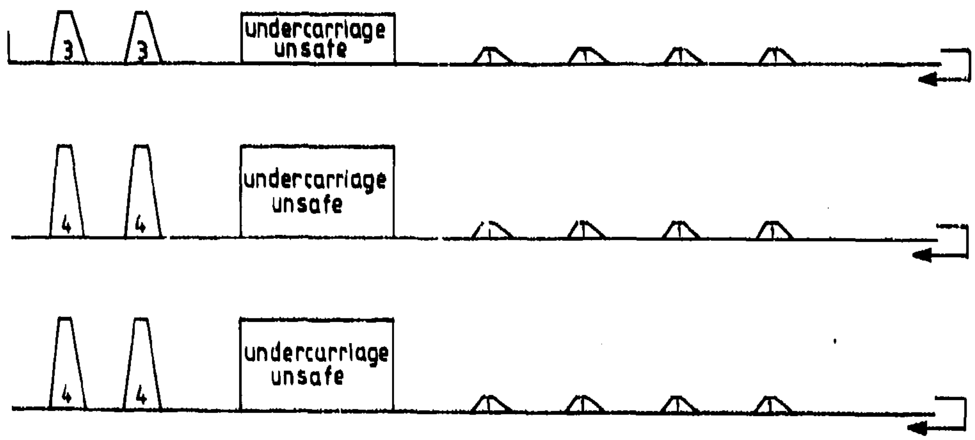

Figure 10. Time course of a complete auditory warning signal with voice message.

Note. From Guidelines for Auditory Warning Systems on Civil Aircraft (CAA Paper 82017) by R. D. Patterson, 1982, London: Civil Aviation Authority. Reprinted by permission.

To assist in tailoring warning signal parameters to specific aircraft, Lower and Wheeler (1985) have developed a desk-top computer program to predict masked thresholds for warning signals in a given noise environment. The program has been validated by comparing measured and predicted thresholds in recorded noise from Chinook, Sea King, and Lynx helicopters. The investigators found a high correlation between measured and predicted masked thresholds (Lower and Wheeler, 1985).

Coleman and his colleagues at the U.K.'s National Coal Board drew heavily on Patterson's work in developing guidelines for warning signals in industrial operations in general and coal production in particular (Coleman et al., 1984). They noted that Patterson's guidelines were designed for aircraft cockpits, but they could be effectively used in certain other environments such as control rooms. They also pointed out that Patterson's signals differed mainly in the temporal domain, whereas frequency and level characteristics could also be varied. In addition to Patterson's technique, Coleman et al. relied on ideas from Deatherage (1972) and Licklider (1961) in formulating the following recommendations:

Guidelines for the Production of Discriminable Sets of Signals (from Coleman et al., 1984, pp. 60-61)

1. Limit the number of signals to six at any one workplace.

2. Use no more than two signals when only one signal characteristic, such as pitch, is altered.

3. Ensure at least three harmonically, or pseudo-harmonically related spectral components occur in the range 1-2 kHz for each signal.

4. Ensure that signals differ both in terms of their temporal patterns and their constituent perceptual units.

5. To manipulate temporal pattern use modulation (AM or FM) at rates of 1 to 4 Hz, employing rest periods between bursts of sound as part of the temporal pattern.

6. Ensure that the modulation rate does not correspond with fluctuation rates in the environmental noise.

7. To manipulate within perceptual units, use different pitch, and higher frequency modulation (AM or FM) at rates above 20 Hz. To manipulate pitch it is best to use complex signals comprising several harmonically related components. Such signals have a fixed perceived pitch regardless of the particular order of the harmonics (see Plomp, 1967). Masking some of the components will not alter the perceived pitch and signals made up of many harmonically related components can, therefore, be more resistant to the effects of short term noises, both in terms of maintaining their audibility and perceived identity. In following this recommendation it should be remembered that the frequency of the fundamental present in the signal or implied by the harmonics should be below 1 kHz.

VIII. SUMMARY

Speech Variables

The proper assessment of speech communication conditions requires knowledge of the speech level. A number of different methods of measuring speech level are in use now, yielding differing results, although the relationships among these methods are fairly stable. People do not always talk at the same level. In live-voice situations, it should be kept in mind that people raise their voices about 5-6 dB for every 10-dB increase in background noise, and that vocal effort increases with distance and even with different forms of activity.

Normal speech is highly redundant, especially the special phraseologies that are often used in military situations. Because of conditions of noise and filtering, however, redundancy is significantly reduced. There are a variety of useful speech materials available, and it is important to select appropriate materials to evaluate the particular noise conditions, communication system, and communication needs at hand.

We seldom listen to speech in ideal circumstances. Filtering characterizes communication systems, and noise masking is the most common source of speech interference. The upward spread of masking makes high levels of noise disproportionately disruptive. Combined distortions act synergistically to degrade communication.

Transmission Characteristics

Intelligibility of the speech signal is modified by distance, reverberation, and spatial location with respect to the noise source. Speech level is reduced by 6 dB per doubling of distance outdoors, but the reduction is less indoors because of reverberant build-up. Reverberation begins to degrade speech intelligibility at about 0.8 second in quiet, and at less than 0.5 second in noisy backgrounds. The effect is greater in small than it is in large rooms. Separation in space of the speech and noise signals can result in improvements equivalent to a speech-to-noise ratio of 5 dB.

Binaural listening provides improvements of anywhere from 2.5 to 13 dB, depending mainly on noise and reverberation conditions. Telephone listening is difficult in noise because the filtering involved reduces speech redundancy, and background noise reduces it further. Intelligibility can be improved by reducing side-tone feedback, through occluding or modifying the transmitter or by amplifying the signal. Current communication systems are frequently outmoded, causing strain and delays on the part of the listener, but there are many possibilities for improvement.

Talker and Listener Variables

Although individuals are capable of producing voice levels as high as 100-105 dB, they cannot sustain speaking levels above an asymptotic level of about 78 dB without considerable discomfort. Individuals who must habitually communicate in noise over a period of years are subject to voice disorders, such as hoarseness and vocal nodules. Talker articulation can greatly affect speech intelligibility. Studies have shown improvements of up to 18% from speaking clearly in quiet. Improvements also occur in noisy conditions, but appear to be somewhat less dramatic. Female voices are as intelligible as male voices in low and moderate noise levels, but may be slightly less intelligible in high noise levels.

Preferred listening levels under earphones are identified as sound pressure levels of 80-85 dB in quiet. With the introduction of noise, preferred levels are somewhat lower. Tolerable listening levels are lower for negative than they are for positive speech-to-noise ratios. Comfortable listening levels should be at a speech-to-noise ratio of at least 5 dB, and preferably above 10 dB. Non-native listeners have significantly more difficulty understanding degraded speech in their second language, even though they may be fluent speakers of that language. High levels of speech and noise can cause auditory fatigue (both central and peripheral, it appears), which reduces speech discrimination both simultaneously and subsequent to the high-level stimulation.

Prediction Methods

The Articulation Index (AI) is a popular and highly respected method of predicting speech intelligibility in noise. It has been modified and improved by the inclusion of corrections for such conditions as reverberation, spread of masking, peak clipping, changes in vocal effort, lipreading, and hearing impairment.

Speech Interference Level (SIL) is useful for predicting distances at which "just reliable" communication can occur. It has recently been modified to apply to indoor situations and numerous levels of vocal effort. But its utility is limited because it cannot be used for hearing-impaired people or for other than face-to-face communication situations, and it uses only one level of intelligibility.

The Speech Transmission Index (STI) takes into account volume and reverberation time of the room, noise level, and distance between talker and listener, yielding a value similar to the AI. Research by the Netherlands group that developed the STI shows this method to be a good predictor of speech communication in a wide variety of conditions.

The sound level meter's A-weighting network can be a good predictor of speech interference, and has the advantage of being inexpensive and easy to use. Other interesting weighting networks have been proposed, but have not been incorporated into the standard sound level meter.

Although the above schemes use different measurement methods, the products can be related in a fairly predictable way. For example, 70% monosyllable intelligibility can be achieved at an AI of 0.45 or an STI of 0.55, which corresponds to a speech-to-noise ratio of about 1.5 dB.

Acceptability Criteria

There is general agreement that minimal or "just reliable" communication can take place at an AI of 0.3 to 0.45, but those who recommend these values give little information about the use of this level of communication. Although the literature is virtually silent on the specific types and amounts of communication needed for various operations, there are numerous recommendations for the range of background noise levels appropriate for certain activities and spaces. Examples include levels of 56-66 dB(A) for shops, garages, and power plant control rooms, down to 21-30 dB(A) for large auditoriums and concert halls.

The only references in the literature to the consequences of degraded speech tend to be anecdotal or subjective. In the military, however, it should be obvious that normal patterns of communication can break down in emergencies, and the consequences of misunderstood instructions can be as serious as destruction of expensive property, or even loss of life.

Detection of Warning Signals

Because a warning signal is detectable does not necessarily mean it will be effective. Ideally, a signal should be at least 15 dB but no more than 25 dB above its masked threshold. Temporal, spectral, and ergonomic aspects should emphasize attention demand, relevance, and appropriate level of priority, without being unduly aversive.

IX. RESEARCH RECOMMENDATIONS

1. Probably the most important information about the effects of noise on military speech communication would be an assessment of the consequences of communication failures. Because servicemen, and especially aircrews, have "admirable tendencies to refrain from complaining" (see Money, 1981), this information is not easily gained. Perhaps the best approach would be a carefully worded survey of personnel working in high-noise environments, such as tanks and helicopters, which would promise anonymity.

2. Another difficult but important project is to assess the type, amount, spectrum, dynamic range, actual content, and intelligibility of speech communication needed for the most efficient conduct of specific tasks. Once these factors are known, communication systems can be successfully matched to the various tasks.

3. A worthwhile research area that has received very little attention in this country is the effect of high levels of vocal effort on the larynx and the incidence of vocal abnormalities and pathologies among personnel who communicate in high noise levels. A related topic would be the influence of vocal strain and intense vibration (as in helicopters and tanks) on speech intelligibility.

4. A final area would be an assessment of the adequacy of auditory warning signals in view of the research and guidelines of Patterson and his colleagues.

REFERENCES

Allen, J.B. and Berkley, D.A. Image method for efficiently simulating small-room acoustics. J. Acoust. Soc. Am., 65, 943-950, 1979.

Anon. Noisy work situations may impair vocal cords. J. Amer. Med. Assoc., 241, 2133, 1979.

ANSI. American National Standard Method for Measurement of Monosyllabic Word Intelligibility. ANSI S3.2-1960 (R1982).

ANSI. American National Standard: Methods for the Calculation of the Articulation Index. ANSI S3.5, 1969.

ASA. Acoustical Society of America. Proposed American National Standard Method for Measuring the Intelligibility of Speech Over Communication Systems. ANSI S3.2-198X. (First draft, July 1988).

ASA. Acoustical Society of America. American National Standard Method for Determining the Electrical Level of Speech. ANSI S3.38, 198X. (Draft dated September 1986).

ASA. Acoustical Society of America. American National Standard for Rating Noise with Respect to Speech Interference. ANSI S3.14, 1977.

Beattie, R.C. Comfortable speech listening levels: Effects of competition and instructions. Am. J. Otol., 3, 192-199, 1982.

Beattie, R.C. and Himes, B.A. Comfortable loudness levels for speech: Effects of signal-to-noise ratios and instructions. J. Aud. Res., 24, 213-229, 1984.

Beranek, L.L. Revised criteria for noise in buildings. Noise Control., 3, 19-27, 1957.

Beranek, L.L. Acoustics. N.Y.: McGraw-Hill, 1954.

Beranek, L.L. Noise control in office and factory spaces. Ind. Hyg. Found. Am., Trans. Bull. 18, 26-33, 1950.

Beranek, L.L. Blazier, W.E., and Figwer, J.J. Preferred noise criterion (PNC) curves and their application to rooms. J. Acoust. Soc. Am., 50, 1223-1228, 1971.

Bilger, R.C. Speech recognition test development. In E. Elkins (Ed.) Speech Recognition by the Hearing Impaired. ASHA Reports, 14, 1984.

Black, J.W. Reception of repeated and overlapping speech patterns. J. Acoust. Soc. Am., 27, 494, 1955.

Black, J.W. Accompaniments of word intelligibility. J. Speech Hear. Disord., 17, 409-418, 1952.

Borchgrevink, H.M. Second language speech comprehension in noise - A hazard to aviation safety. In K.E. Money, (Ed.) Aural Communications in Aviation. NATO, Advisory Group for Aerospace Research Development. AGARD-CP-311, 1981.

Bradley, J.S. Predictors of speech intelligibility in rooms. J. Acoust. Soc. Am., 80, 837-845, 1986.

Brady, P.T. Equivalent peak level: A threshold-independent speech-level measure. J. Acoust. Soc. Am., 44, 695-699, 1968.

Braida, L.D., Durlach, N.I., Lippmann, R.P., Hicks, B.L., Rabinowitz, W.R., and Reed, C.M. Hearing aids: A review of past research on linear amplification, amplitude compression, and frequency lowering. Amer. Speech & Hearing Assoc., Monogr. 19, 1979.

Brandy, W.T. Reliability of voice tests of speech discrimination. J. Speech Hear. Res., 9, 461-465, 1966.

Broadbent, D.E. Perception and Communication. London, Pergamon, 1958.

CHABA. Committee on Hearing, Bioacoustics, and Biomechanics. Guidelines for Noise and Vibration Levels for the Space Station. NASA Contractor Report 178310, National Aeronautics and Space Administration, Washington, D.C., 1987.

CHABA. Committee on Hearing, Bioacoustics, and Biomechanics. The Effects of Time-Varying Noise on Speech Intelligibility Indoors. Report of Working Group 83. Washington, D.C., National Academy Press, 1981.

CHABA. Committee on Hearing, Bioacoustics, and Biomechanics. A proposed standard fire alarm signal. J. Acoust. Soc. Am., 57, 756-757, 1975.

Clarke, F.R. Technique for evaluation of speech systems. Stanford Research Institute Project 5090. U.S. Army Electronics Laboratory, Contract DA 28-043 AMC-00227 (E), 1965.

Coleman, G.J., Graves, R.J., Collier, S.G., Golding, D., Nicholl, A.G. McK., Simpson, G.C., Sweetland, K.F., and Talbot, C.F. Communications in noisy environments. Final Report on CEC Contract 7206/00/8/09. Institute of Occup. Med., National Coal Board, Burton-on-Trent, U.K., 1984.

Deatherage, B.H. Auditory and other sensory forms of information presentation. In: H.P. Van Cott and R.G. Kinkade (Eds.), Human Engineering Guide to Equipment Design. Wash., D.C., U.S. Govt. Printing Office, 1972.

DoD. Department of Defense. Human engineering design criteria for military systems, equipment and facilities. MIL-STD-1472C. Washington, D.C., 1981.

Dugal, R.L., Braida, L.D., and Durlach, N.I. Implications of previous research for the selection of frequency-gain characteristics. In G. Studebaker and I. Hochberg (Eds.) Acoustical Factors Affecting Hearing Aid Performance. Baltimore: Univ. Park Press, 1980.

EPA. U.S. Environmental Protection Agency. Information on levels of environmental noise requisite to protect public health and welfare with an adequate margin of safety. EPA 550/9-74-004. Washington, D.C., 1974.

Fairbanks, G. Test of phonemic differentiation: The Rhyme Test. J. Acoust. Soc. Am., 30, 596-600, 1958.

Florentine, M. Speech perception in noise by fluent, non-native listeners. J. Acoust. Soc. Am., Suppl. 1, 77, S106, 1985.

French, N.R. and Steinberg, J.C. Factors governing the intelligibility of speech sounds. J. Acoust. Soc. Am., 19, 90-119, 1947.

Frick, F.C. and Sumby, W.H. Control tower language. J. Acoust. Soc. Am., 24, 595-596, 1952.

Gardner, M.B. Effect of noise, system gain, and assigned task on talking levels in loudspeaker communication. J. Acoust. Soc. Am., 40, 955-965, 1966.

Garinther, G.R. and Hodge, D.C. Effect of the M25 protective mask and hood on speech intelligibility and voice level. U.S. Army Human Engineering Lab., TM 19-87, Aberdeen Proving Ground, MD, 1987.

Harris, J.D. Pure-tone acuity and the intelligibility of everyday speech. J. Acoust. Soc. Am., 37, 824-830, 1965.

Harris, J.D. Combinations of distortions in speech: The twenty-five percent safety factor by multiple-cueing. Arch. Otolaryngol., 72, 227-232, 1960.

Harris, J.D., Haines, H.L., Kelsey, P.A., and Clack, T.D. The relation between speech intelligibility and electroacoustic characteristics of low fidelity circuitry. J. Aud. Res., 1, 357-381, 1961.

van Heusden, E., Plomp, R., and Pols, L.C.W. Effect of ambient noise on the vocal output and the preferred listening level of conversational speech. Applied Acoustics, 12, 31-43, 1979.

Hirsh, I.J., Davis, H., Silverman, S.R., Reynolds, E.G., Eldert, E., and Benson, R.W. Development of materials for speech audiometry. J. Speech Hear. Disord. 17, 321-337, 1952.

Holmes, A.E., Frank, T., and Stoker, R.G. Telephone listening ability in a noisy background. Ear and Hearing, 4, 88-90, 1983.

Hood, J.D. and Poole, J.P. Improving the reliability of speech audiometry. Brit. J. of Audiol., 11, 93-101, 1977.

Hormann, H. and Ortscheid, J. Einfluss von Larm auf Lernvorgange verschiedener zeitlicher Erstreckung. Forschungsbericht 81-1050 104/01 des Umweltbundesamtes, Berlin, 1981.

House, A.S., Williams, C.E., Hecker, M.H.L., and Kryter, K.D. Articulation testing methods: Consonantal differentiation with a closed-response set. J. Acoust. Soc. Am., 37, 158-166, 1965.

Houtgast, T. and Steeneken, H.J.M. Experimental verification of the STI: A valid basis for setting indoor noise level criteria. In G. Rossi (Ed.) Proceedings of the Fourth International Congress on Noise as a Public Health Problem. Centro Ricerche e Studi Amplifon, Milan, Italy, 1983.

Houtgast, T. Indoor speech intelligibility and indoor noise level criteria. In J.V. Tobias, G. Jansen, and W.D. Ward (Eds.) Proceedings of the Third International Congress on Noise as a Public Health Problem. ASHA Reports 10, American Speech-Language-Hearing Assoc., Rockville, MD, 1980.

Hudgins, C.V., Hawkins, J.E., Karlin, J.E., and Stevens, S.S. The development of recorded auditory tests for measuring hearing loss for speech. Laryngoscope, 57, 57-89, 1947.

Humes, L.E., Boney, S., and Loven, F. Further validation of the Speech Transmission Index. J. Speech Hear. Res., 30, 403-410, 1987.

Humes, L.E., Dirks, D.D., Bell, T.S., Ahlstrom, C., and Kincaid, G.E. Application of the Articulation Index and the Speech Transmission Index to the recognition of speech by normal-hearing and hearing-impaired listeners. J. Speech Hear. Res., 29, 447-462, 1986.

Jerger, J., Speaks, C., and Trammell, J.L. A new approach to speech audiometry. J. Speech Hear. Disord., 33, 318-328, 1968.

Jones, D. and Broadbent, D. Side-effects of interference with speech by noise. Ergonomics, 22, 1073-1081, 1979.

Kalikow, D.N., Stevens, K.N., and Elliott, L.L. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. J. Acoust. Soc. Am., 61, 1337-1351, 1977.

Kamm, C.A., Dirks, D.D., and Bell, T.S. Speech recognition and the Articulation Index for normal and hearing-impaired listeners. J. Acoust. Soc. Am., 77, 281-288, 1985.

Klingholz, F., Siegert, C., Schleier, E., and Thamm, R. Larmbedingte Stimmstorungen bei Angehorigen unterschiedlicher Berufsgruppen. HNO-Praxis, 3/3, 193-201, 1978.

Klumpp, R.G. and Webster, J.C. Physical measurements of equally speech interfering navy noises. J. Acoust. Soc. Am., 35, 1328-1338, 1963.

Klumpp, R.G. and Webster, J.C. USNEL flight deck communications system: Part 5, message analysis and personnel response. NEL Rep. 926, 1960.

Kreul, E.J., Nixon, J.C., Kryter, K.D., Bell, D.W., Lang, J.S., and Schubert, E.D. A proposed clinical test of speech discrimination. J. Speech Hear. Res., 11, 536-552, 1968.

Kryter, K.D. Physiological, Psychological, and Social Effects of Noise. NASA Reference Publication 1115, National Aeronautics and Space Administration, Washington, D.C., 1984.

Kryter, K.D. The Effects of Noise on Man. New York: Academic Press, 1970.

Kryter, K.D. Methods for the calculation and use of the Articulation Index. J. Acoust. Soc. Am., 34, 1689-1697, 1962a.

Kryter, K.D. Validation of the Articulation Index. J. Acoust. Soc. Am., 34, 1698-1702, 1962b..

Kryter, K.D. Effects of ear protective devices on the intelligibility of speech in noise. J. Acoust. Soc. Am., 18, 413-417, 1946.

Kryter, K.D. and Williams, C.E. Some factors influencing human response to aircraft noise: Masking of speech and variability of subjective judgements. FAA-ADS-42, June 1965.

Kuttruff, H. Room Acoustics. London: Applied Science Publishers, 1973.

Lacroix, P.G., Harris, J.D., and Randolph, K.J. Multiplicative effects on sentence comprehension for combined acoustic distortions. J. Speech Hear. Res., 22, 259-269, 1979.

Lazarus, H. Literature reviews 1978-1983 in German language on noise and communication. In G. Rossi (Ed.) Proceedings of the Fourth International Congress on Noise as a Public Health Problem. Centro Ricerche e Studi Amplifon, Milan, Italy, 1983.

Levitt, H. and Rabiner, L.R. Predicting binaural gain in intelligibility and release from masking for speech. J. Acoust. Soc. Am., 42, 820-829, 1967.

Licklider, J.C.R. Audio warning signals for Air Force weapon systems, WADD TR-60-814, Dayton, OH, Wright-Patterson AFB, 1961.

Licklider, J.C.R. and Pollack, I. Effects of differentiation, integration and infinite peak clipping upon the intelligibility of speech. J. Acoust. Soc. Am., 20, 42-51, 1948.

Littler, T.S. Physics of the Ear. London: Pergamon, 1965.

Lochner, J.P.A. and Burger, J.F. The influence of reflections on auditorium acoustics. J. Sound Vib., 1, 426-454, 1964.

Lower, M.C. and Wheeler, P.D. Specifying the sound levels for auditory warnings in noisy environments. Paper presented at the 9th Congress of the International Ergonomics Association. Bournemouth, England, 1985.

Loftiss, E. An evaluation of the effects of selected psychophysical methods and judgmental procedures upon comfortable loudness judgments of connected speech. Unpubl. doctoral diss., Univ. Illinois, 1964.

MacKeith, N.W. and Coles, R.R.A. Binaural advantages in hearing of speech. _J. Laryng. Otol._, _85_, 213-232, 1971.

Mankovsky, V.S. _Acoustics of Studios and Auditoria._ C. Gilford (Ed.), A.G. Clough (Trans.). London: Focal Press, 1971.

Martin, D.W., Murphy, R.L. and Meyer, A. Articulation reduction of combined distortions of speech waves. _J. Acoust. Soc. Am._, _28_, 597-601, 1956.

Mayer, M.S. and Lindburg, A.W. Electronic voice communications improvements for Army aircraft. In K.E. Money (Ed.) _Aural Communication in Aviation_, NATO, Advisory Group for Aerospace Research Development. AGARD-CP-311, 1981.

McKinley, R.L. Voice communication research and evaluation system. In K.E. Money (Ed.) _Aural Communications in Aviation_. NATO, Advisory Group for Aerospace Research Development. AGARD-CP-311, 1981.

Miller, G.A. The masking of speech. _Psych. Bull._, _44_, 105-129, 1947.

Miller, G.A., Heise, G.A., and Lichten, W. The intelligibility of speech as a function of the context of the test materials. _J. Exp. Psych._, _41_, 329-335, 1951.

Money, K.E. Technical evaluation report. In K.E. Money (Ed.) _Aural Communications in Aviation_. NATO, Advisory Group for Aerospace Research Development. AGARD-CP-311, 1981.

Moore, T.J., Nixon, C.W., and McKinley, R.L. Comparative intelligibility of speech materials processed by standard Air Force voice communication systems in the presence of simulated cockpit noise. In K.E. Money (Ed.) _Aural Communications in Aviation_, NATO, Advisory Group for Aerospace Research Development. AGARD-CP-311, 1981.

Moser, H.M. and Dreher, J.J. Phonemic confusion vectors. _J. Acoust. Soc. Am._, _27_, 874-881, 1955.

Moser, H.M. and Dreher, J.J. Effects of training on listeners in intelligibility studies. _J. Acoust. Soc. Am._, _27_, 1213-1219, 1955.

Moser, H.M., Dreher, J.J., and Adler, R.F. Project 519, Rept. No. 11, AFCRC-TR-54-81, 1954.

Mosko, J.D. "Clear speech." A strategem for improving radio communications and automatic speech recognition in noise. In K.E. Money (Ed.) _Aural Communications in Aviation_, NATO, Advisory Group for Aerospace Research Development. AGARD-CP-311, 1981.

Nabelek, A.K. Acoustics of enclosed spaces for linguistically impaired listeners. In G. Rossi (Ed.) _Noise as a Public Health Problem_: _Proceedings of the Fourth International Congress_. Centro Ricerche e Studi Amplifon, Milan, Italy, 1983.

Nabelek, A. Distortions and age effect on speech in noise. In J.V. Tobias, G. Jansen, and W.D. Ward (Eds.) Proceedings of the Third International Congress on Noise as a Public Health Problem: ASHA Reports 10, American Speech-Language-Hearing Assoc., Rockville, MD, 1980.

Nabelek, A.K. and Pickett, J.M. Reception of consonants in a classroom as affected by monaural and binaural listening, noise, reverberation, and hearing aids. J. Acoust. Soc. Am., 56, 628-639, 1974.

Nabelek, A.K. and Robinette, L. Reverberation as a parameter in clinical testing. Audiology, 17, 239-259, 1978.

NAVSHIPS. Shipboard Aviation Maintenance Communication, 0900-072-2010, 1972.

Parker, D.E., Martens, W.L., and Johnston, P.A. Influence of auditory fatigue on masked speech intelligibility. J. Acoust. Soc. Am., 67, 1392-1393, 1980.

Patterson, R.D. Auditory warning systems for high-workload environments. Paper presented at the 9th Congress of the International Ergonomics Assoc., Bournemouth, England, 1985.

Patterson, R.D. Guidelines for Auditory Warning Systems on Civil Aircraft, CAA Paper 82017, Civil Aviation Authority, London, 1982.

Pearsons, K.S. Standardized methods for measuring speech levels. In G. Rossi (Ed.) Proceedings of the Fourth International Congress on Noise as a Public Health Problem. Centro Ricerche e Studi Amplifon, Milan, Italy, 1983.

Pearsons, K.S. Communication in noise: Research after the 1973 Congress on Noise as a Public Health Problem. In J.V. Tobias, G.J. Jansen, and W.D. Ward (Eds.) Proceedings of the Third International Congress on Noise as a Public Health Problem. ASHA Reports 10, The American Speech-Language-Hearing Association, Rockville, Maryland, 1980.

Pearsons, K.S., Bennett, R.L., and Fidell, S. Speech Levels in Various Noise Environments. EPA-600/1-77-025. U.S. Environmental Protection Agency, 1977.

Peterson, G.E. and Lehiste, I. Revised CNC lists for auditory tests. J. Speech Hear. Disord., 27, 62-70, 1962.

Picheny, M.A., Durlach, N.I., and Braida, L.D. Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech. J. Sp. and Hear. Res., 29, 434-446, 1986.

Picheny, M.A., Durlach, N.I., and Braida, L.D. Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech. J. Sp. and Hear. Res., 28, 96-103, 1985.

Pickett, J.M. Effects of vocal force on the intelligibility of speech sounds. J. Acoust. Soc. Am., 28, 902-905, 1956.

Platte, H. J. Kouzeption eines "sinnvollen" Sprachverstandnistests unter Storschalleinfluss. *Z. Horgerate-Akustik*, *18*, 208-222, 1979.

Platte, H. J. Probleme bei Sprachverstandnistests unter Storschalleinfluss. *Z. Horgerate-Akustik*, *17*, 190-207, 1978.

Plomp, R. Binaural and monaural speech intelligibility of connected discourse in reverberation as a function of azimuth of a single competing sound source (speech or noise). *Acustica*, *34*, 200-210, 1976.

Plomp, R. Pitch of complex tones. *J. Acoust. Soc. Am.*, *41*, 1526-1533, 1967.

Pollack, I. Speech intelligibility at high noise levels: Effect of short-term exposure. *J. Acoust. Soc. Am.*, *30*, 282-285, 1958.

Pollack, I. and Pickett, J.M. Masking of speech by noise at high sound levels. *J. Acoust. Soc. Am.*, *30*, 127-130, 1958.

Pols, L.C.W., van Heusden, E., and Plomp, R. Preferred listening level for speech disturbed by fluctuating noise. *Applied Acoustics*, *13*, 267-279, 1980.

Rabbitt, P.M.A. Channel capacity intelligibility and immediate memory. *Quart. J. Expt. Psych.*, *20*, 241-248, 1968.

Rabbitt, P.M.A. Recognition: Memory for words heard correctly in noise. *Psychonomic Science*, *8*, 383-384, 1966.

Richards, D.L. *Telecommunication by Speech: The Transmission Performance of Telephone Networks*. London: Butterworths, 1973.

Rood, G.M., Chillary, J.A., and Collistor, J.B. Requirements and application of auditory warnings to military helicopters. Paper presented at the 9th Congress of the International Ergonomics Association. Bournemouth, England, 1985.

Rubenstein, H. and Pollack, I. Word predictability and intelligibility. *J. Verb. Learn. Verb. Behav.*, *2*, 147-158, 1963.

Rubenstein, H., Decker, L., and Pollack, I. Word length and intelligibility. *Lang. Speech*, *2*, 175-177, 1959.

Rupf, J.A. Noise effects on passenger communication in light aircraft. Paper No. 770446, Soc. Automot. Eng., Warrendale, PA, 1977.

Schaenman, H. Factors influencing judgment of the level of most comfortable loudness (MCL) for speech. Unpub. M.A. thesis, Brooklyn College, 1965.

Schleier, E. Klinische und stroboskopische Befunde bei Larmarbeitern. *Dt. Gesundh.-Wesen*, *32*, 123-126, 1977.

Schultz, T.J. Predicting sound pressure levels in furnished dwellings and offices. *J. Acoust. Soc. Am.*, *Suppl. 1*, *75*, S90, 1984.

Silverman, S.R. and Hirsh, I.J. Problems related to the use of speech in clinical audiometry. Ann. Otol. Rhinol. Laryngol, 64, 1234-1244, 1955.

Skinner, M.W. and Miller, J.D. Amplification bandwidth and intelligibility of speech in quiet and noise for listeners with sensorineural hearing loss. Audiology, 22, 253-279, 1983.

Snidecor, J.C., Malbry, L.A., and Hearsey, E.L. Methods of training talkers for increasing intelligibility in noise. ORSD Report 3178. New York: The Psychological Corp., 1944.

Sorin, C. and Thouin-Daniel, C. Effects of auditory fatigue on speech intelligibility and logical decision in noise. J. Acoust. Soc. Am., 74, 456-466, 1983.

Speaks, C. and Jerger, J. Method for measurement of speech identification. J. Speech Hear. Res., 8, 185-194, 1965.

Steeneken, H.J.M. and Houtgast, T. Comparison of some methods for measuring speech levels. Institute for Perception TNO Report IZF 1978-22. Soesterberg, Netherlands, 1978.

Suter, A.H. The ability of mildly hearing-impaired individuals to discriminate speech in noise. U.S. EPA and U.S. Air Force joint report: EPA 550/9-78-100 and AMRL-TR-78-4, Washington, D.C., 1978.

Swets, J.A., Green, D.M., Fay, T.H., Kryter, K.D., Nixon, C.M., Riney, J.S., Schultz, T.J., Tanner, W.P. and Whitcomb, M.A. A proposed standard fire-alarm signal. J. Acoust. Soc. Am., 57, 756-757, 1975.

Thorndike, E.L. and Lorge, I. The Teacher's Word Book of 30,000 Words. 2nd Ed. New York: Columbia Univ. Press, 1952.

Thwing, E.J. Effect of repetition on articulated scores for PB words. J. Acoust. Soc. Am., 28, 302-303, 1956.

Tillman, T.W. and Carhart, R. An expanded test for speech discrimination utilizing CNC mon('labic words. Northwestern University Auditory Test No. 6. Tech. Rep. SAM-TR-66-55, U.S.A.F. School of Aerospace Medicine, Aerospace Med., Brooks AFB, TX, 1966.

Tolhurst, G.C. Effects of duration and articulation changes on intelligibility, word reception, and listener preference. J. Speech Hear. Disord., 22, 328, 1957.

Tolhurst, G.C. The effect on intelligibility scores of specific instructions regarding talking. USAM Report NM 001 064 01 35. U.S. Naval Air Station, Pensacola, FL, 1955.

U.S. Army Test and Eval. Command. Operational intelligibility testing of voice communication equipment. Aberdeen Proving Ground MTP-6-3-521, AD 718774, 1971.

Voiers, W.D. Diagnostic evaluation of speech intelligibility in M.E. Hawley (Ed.) Speech Intelligibility and Recognition, Dowden, Hutchinson, & Ross, Stroudsburg, PA., 1977.

Webster, J.C. Noise and communication. In D.M. Jones and A.J. Chapman (Eds.) Noise and Society. London, John Wiley & Sons, 1984.

Webster, J.C. Communicating in noise, 1978-1983. In G. Rossi (Ed.) Noise as a Public Health Problem. Proceedings of the Fourth International Congress. Centro Ricerche e Studi Amplifon, Milan, Italy, 1983.

Webster, J.C. Speech interference aspects of noise. In D.M. Lipscomb (Ed.) Noise and Audiology. Baltimore, Univ. Park Press, 1978.

Webster, J.C. Compendium of speech testing material and typical noise spectra to be used in evaluating communication equipment. Tech. Doc. 191. Human Factors Div., Naval Electronics Laboratory Center, San Diego, CA, 1972.

Webster, J.C. Relations between Speech-Interference contours and idealized Articulation-Index contours. J. Acoust. Soc. Am., 36, 1662-1669, 1964.

Webster, J.C. and Allen, C.R. Speech intelligibility in naval aircraft radios. TR 1830. Nav. Electronics Lab. Cent., San Diego, CA, 1972.

Webster, J.C. and Klumpp, R.G. Effects of ambient noise and nearby talkers on a face-to-face communication task. J. Acoust. Soc. Am., 34, 936-941, 1962.

Wilkins, P.A. and Martin, A.M. The effect of hearing protectors on the perception of warning sounds. In P.W. Alberti (Ed.) Personal Hearing Protection in Industry. New York: Raven Press, 1982.

Williams, C.E., Mosko, J.D., and Greene, J.W. Evaluating the ability of air crew personnel to hear speech in their operational environments. Aviation, Space, & Environ. Med., 47, 154-158, 1976.

Williams, C.E., Forstall, J.R., and Parsons, W.C. The effect of earplugs on passenger speech reception in rotary-wing aircraft. NAMRL Rep. 1121. U.S. Naval Aerospace Med. Cent., Pensacola, FL, 1970.